humanVR1 gene with translation of open reading frame

Input file Fchrb87a6.seq; Output File Fchrb87a6.tra
Sequence length 3909

GTGAGCGCAACGCACTGCGGGCAGTGAGCGCAACGCACTGCGGGCAGTGAGCGCAACGCACTGCGGGCAGTGAGCGCAA

CGCACTGCGGGCAGTGAGCGCAACGCACTGCGGGCAGTGAGCGCAACGCACTGCGGGCAGTGAGCGCAACGCACTGCGG

GCAGTGAGCGCAACGCACTTGCGGGCAGTGAGCGCAACGCACTGCGGGCAGTGAGCGCAACGCACTGCGGGCAGTGAGC

GCAACGCACTGCGGGCAGTGAGCGCAACGCACTGCGGGCAGTGAGCGCAACGCACTGCGGGCAGTGAGCGCAACGCACT

GCGGGCAGTGAGCGCAACGCACTGCGGGCAGTGAGCGCAACGCACTGCGGGCAGTGAGCGCAACGCACTGCGGGCAGTG

AGCGCAACGCACTGCGGGCAGTGAGCGCAACGCACTTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACAC

TTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATT

ACGCCAAGCTCTAATACGACTCACTATAGGGAAAGCTGGTACGCCTGCAGGTACCGGTCCGGAATTCCCGGGTCGACCC

ACGCGTCCGAAAACACACCTCTCTGCTGTGGGAAGACTGTGCAATGGCACAGCCGCAGAGCTTGGTTTGGGAGGTTGAA

GTGCTCTGGGGAGAATTCGTAGATCATCCTCAGAAAAGCCTTGCCCTGGTGTTCTACCAGAAAAACGTCTCCCAATCAC

CCAGAAAAGCTGTCCACAGTAGTCCCCCCCTTATCCACGGGTGTCACTTTCCATGGGTTCAGTTATTTGCGGTCAACCAC

GGTCTGCCAATATTAAATGGAAAATTCTTCAAACAGTTCCCAAGTTTTCCCTTGTGCATTGTTCTGAGCAGTGTGATGA

AGAGTCTCTGCCGTGCCATCTGGGATGCAAACCGTCCCTGTGTCCCCCACGTCCAGGCCGTAGATGCTCCCCGCCGGTC

AGTCACTTAGTCGTCAGATCGCCCGTCCTGGTATCACAGTGCTTCTGTTCAGGTTGCACACTGGGCCACAGAGGATCCA

```
         M   K   K   W   S   S   T   D   L   G   T   A   A   D   P   L   Q   K    18
GCAAGG ATG AAG AAA TGG AGC AGC ACA GAC TTG GGG ACA GCT GCG GAC CCA CTC CAA AAG    54

 D   T   C   P   D   P   L   D   G   D   P   N   S   R   P   P   P   A   K   P    38
GAC ACC TGC CCA GAC CCC CTG GAT GGA GAC CCT AAC TCC AGG CCA CCT CCA GCC AAG CCC   114

 Q   L   P   T   A   K   S   R   T   R   L   F   G   K   G   D   S   E   E   A    58
CAG CTC CCC ACG GCC AAG AGC CGC ACC CGG CTC TTT GGG AAG GGT GAC TCG GAG GAG GCT   174

 F   P   V   D   C   P   H   E   E   G   E   L   D   S   C   P   T   I   T   V    78
TTC CCG GTG GAT TGC CCC CAC GAG GAA GGT GAG TTG GAC TCC TGC CCG ACC ATC ACA GTC   234

 S   P   V   I   T   I   Q   R   P   G   D   G   P   T   G   A   R   L   L   S    98
AGC CCT GTT ATC ACC ATC CAG AGG CCA GGA GAC GGC CCC ACC GGT GCC AGG CTG CTG TCC   294

 Q   D   S   V   A   A   S   T   E   K   T   L   R   L   Y   D   R   R   S   I   118
CAG GAC TCT GTC GCC GCC AGC ACC GAG AAG ACC CTC AGG CTC TAT GAT CGC AGG AGT ATC   354

 F   E   A   V   A   Q   N   N   C   Q   D   L   E   S   L   L   L   F   L   Q   138
TT GAA GCC GTT GCT CAG AAT AAC TGC CAG GAT CTG GAG AGC CTG CTG CTC TTC CTG CAG    414
```

FIGURE 1

```
  K   S   K   K   H   L   T   D   N   E   F   K   D   P   E   T   G   K   T   C    158
 AAG AGC AAG AAG CAC CTC ACA GAC AAC GAG TTC AAA GAC CCT GAG ACA GGG AAG ACC TGT   474


  L   L   K   A   M   L   N   L   H   D   G   Q   N   T   T   I   P   L   L   L    178
 CTG CTG AAA GCC ATG CTC AAC CTG CAC GAC GGA CAG AAC ACC ACC ATC CCC CTG CTC CTG   534


  E   I   A   R   Q   T   D   S   L   K   E   L   V   N   A   S   Y   T   D   S    198
 GAG ATC GCG CGG CAA ACG GAC AGC CTG AAG GAG CTT GTC AAC GCC AGC TAC ACG GAC AGC   594


  Y   Y   K   G   Q   T   A   L   H   I   A   I   E   R   R   N   M   A   L   V    218
 TAC TAC AAG GGC CAG ACA GCA CTG CAC ATC GCC ATC GAG AGA CGC AAC ATG GCC CTG GTG   654


  T   L   L   V   E   N   G   A   D   V   Q   A   A   A   H   G   D   F   F   K    238
 ACC CTC CTG GTG GAG AAC GGA GCA GAC GTC CAG GCT GCG GCC CAT GGG GAC TTC TTT AAG   714


  K   T   K   G   R   P   G   F   Y   F   G   E   L   P   L   S   L   A   A   C    258
 AAA ACC AAA GGG CGG CCT GGA TTC TAC TTC GGT GAA CTG CCC CTG TCC CTG GCC GCG TGC   774


  T   N   Q   L   G   I   V   K   F   L   L   Q   N   S   W   Q   T   A   D   I    278
 ACC AAC CAG CTG GGC ATC GTG AAG TTC CTG CTG CAG AAC TCC TGG CAG ACG GCC GAC ATC   834


  S   A   R   D   S   V   G   N   T   V   L   H   A   L   V   E   V   A   D   N    298
 AGC GCC AGG GAC TCG GTG GGC AAC ACG GTG CTG CAC GCC CTG GTG GAG GTG GCC GAC AAC   894


  T   A   D   N   T   K   F   V   T   S   M   Y   N   E   I   L   M   L   G   A    318
 ACG GCC GAC AAC ACG AAG TTT GTG ACG AGC ATG TAC AAT GAG ATT CTG ATG CTG GGG GCC   954


  K   L   H   P   T   L   K   L   E   E   L   T   N   K   K   G   M   T   P   L    338
 AAA CTG CAC CCG ACG CTG AAG CTG GAG GAG CTC ACC AAC AAG AAG GGA ATG ACG CCG CTG   1014


  A   L   A   A   G   T   G   K   I   G   V   L   A   Y   I   L   Q   R   E   I    358
 GCT CTG GCA GCT GGG ACC GGG AAG ATC GGG GTC TTG GCC TAT ATT CTC CAG CGG GAG ATC   1074


  Q   E   P   E   C   R   H   L   S   R   K   F   T   E   W   A   Y   G   P   V    378
 CAG GAG CCC GAG TGC AGG CAC CTG TCC AGG AAG TTC ACC GAG TGG GCC TAC GGG CCC GTG   1134


  H   S   S   L   Y   D   L   S   C   I   D   T   C   E   K   N   S   V   L   E    398
 CAC TCC TCG CTG TAC GAC CTG TCC TGC ATC GAC ACC TGC GAG AAG AAC TCG GTG CTG GAG   1194


  V   I   A   Y   S   S   S   E   T   P   N   R   H   D   M   L   L   V   E   P    418
 GTG ATC GCC TAC AGC AGC AGC GAG ACC CCT AAT CGC CAC GAC ATG CTC TTG GTG GAG CCG   1254


  L   N   R   L   L   Q   D   K   W   D   R   F   V   K   R   I   F   Y   F   N    438
 CTG AAC CGA CTC CTG CAG GAC AAG TGG GAC AGA TTC GTC AAG CGC ATC TTC TAC TTC AAC   1314


  F   L   V   Y   C   L   Y   M   I   I   F   T   M   A   A   Y   Y   R   P   V    458
 TTC CTG GTC TAC TGC CTG TAC ATG ATC ATC TTC ACC ATG GCT GCC TAC TAC AGG CCC GTG   1374


  D   G   L   P   P   F   K   M   E   K   I   G   D   Y   F   R   V   T   G   E    478
 GAT GGC TTG CCT CCC TTT AAG ATG GAA AAA ATT GGA GAC TAT TTC CGA GTT ACT GGA GAG   1434


  I   L   S   V   L   G   G   V   Y   F   F   F   R   G   I   Q   Y   F   L   Q    498
 ATC CTG TCT GTG TTA GGA GGA GTC TAC TTC TTT TTC CGA GGG ATT CAG TAT TTC CTG CAG   1494
```

FIGURE 1 (cont'd)

R	R	P	S	M	K	T	L	F	V	D	S	Y	S	E	M	L	F	F	L	518
GG	CGG	CCG	TCG	ATG	AAG	ACC	CTG	TTT	GTG	GAC	AGC	TAC	AGT	GAG	ATG	CTT	TTC	TTT	CTG	1554

Q	S	L	F	M	L	A	T	V	V	L	Y	F	S	H	L	K	E	Y	V	538
AG	TCA	CTG	TTC	ATG	CTG	GCC	ACC	GTG	GTG	CTG	TAC	TTC	AGC	CAC	CTC	AAG	GAG	TAT	GTG	1614

A	S	M	V	F	S	L	A	L	G	W	T	N	M	L	Y	Y	T	R	G	558
GCT	TCC	ATG	GTA	TTC	TCC	CTG	GCC	TTG	GGC	TGG	ACC	AAC	ATG	CTC	TAC	TAC	ACC	CGC	GGT	1674

F	Q	Q	M	G	I	Y	A	V	M	I	E	K	M	I	L	R	D	L	C	578
TTC	CAG	CAG	ATG	GGC	ATC	TAT	GCC	GTC	ATG	ATA	GAG	AAG	ATG	ATC	CTG	AGA	GAC	CTG	TGC	1734

R	F	M	F	V	Y	I	V	F	L	F	G	F	S	T	A	V	V	T	L	598
CGT	TTC	ATG	TTT	GTC	TAC	ATC	GTC	TTC	TTG	TTC	GGG	TTT	TCC	ACA	GCG	GTG	GTG	ACG	CTG	1794

I	E	D	G	K	N	D	S	L	P	S	E	S	T	S	H	R	W	R	G	618
ATT	GAA	GAC	GGG	AAG	AAT	GAC	TCC	CTG	CCG	TCT	GAG	TCC	ACG	TCG	CAC	AGG	TGG	CGG	GGG	1854

P	A	C	R	P	P	D	S	S	Y	N	S	L	Y	S	T	C	L	E	L	638
CCT	GCC	TGC	AGG	CCC	CCC	GAT	AGC	TCC	TAC	AAC	AGC	CTG	TAC	TCC	ACC	TGC	CTG	GAG	CTG	1914

F	K	F	T	I	G	M	G	D	L	E	F	T	E	N	Y	D	F	K	A	658
TTC	AAG	TTC	ACC	ATC	GGC	ATG	GGC	GAC	CTG	GAG	TTC	ACT	GAG	AAC	TAT	GAC	TTC	AAG	GCT	1974

V	F	I	I	L	L	L	A	Y	V	I	L	T	Y	I	L	L	L	N	M	678
GTC	TTC	ATC	ATC	CTG	CTG	CTG	GCC	TAT	GTA	ATT	CTC	ACC	TAC	ATC	CTC	CTG	CTC	AAC	ATG	2034

L	I	A	L	M	G	E	T	V	N	K	I	A	Q	E	S	K	N	I	W	698
TTC	ATC	GCC	CTC	ATG	GGT	GAG	ACT	GTC	AAC	AAG	ATC	GCA	CAG	GAG	AGC	AAG	AAC	ATC	TGG	2094

K	L	Q	R	A	I	T	I	L	D	T	E	K	S	F	L	K	C	M	R	718
AAG	CTG	CAG	AGA	GCC	ATC	ACC	ATC	CTG	GAC	ACG	GAG	AAG	AGC	TTC	CTT	AAG	TGC	ATG	AGG	2154

K	A	F	R	S	G	K	L	L	Q	V	G	Y	T	P	D	G	K	D	D	738
AAG	GCC	TTC	CGC	TCA	GGC	AAG	CTG	CTG	CAG	GTG	GGG	TAC	ACA	CCT	GAT	GGC	AAG	GAC	GAC	2214

Y	R	W	C	F	R	V	D	E	V	N	W	T	T	W	N	T	N	V	G	758
AC	CGG	TGG	TGC	TTC	AGG	GTG	GAC	GAG	GTG	AAC	TGG	ACC	ACC	TGG	AAC	ACC	AAC	GTG	GGC	2274

I	I	N	E	D	P	G	N	C	E	G	V	K	R	T	L	S	F	S	L	778
TC	ATC	AAC	GAA	GAC	CCG	GGC	AAC	TGT	GAG	GGC	GTC	AAG	CGC	ACC	CTG	AGC	TTC	TCC	CTG	2334

R	S	S	R	V	S	G	R	H	W	K	N	F	A	L	V	P	L	L	R	798
GG	TCA	AGC	AGA	GTT	TCA	GGC	AGA	CAC	TGG	AAG	AAC	TTT	GCC	CTG	GTC	CCC	CTT	TTA	AGA	2394

E	A	S	A	R	D	R	Q	S	A	Q	P	E	E	V	Y	L	R	Q	F	818
AG	GCA	AGT	GCT	CGA	GAT	AGG	CAG	TCT	GCT	CAG	CCC	GAG	GAA	GTT	TAT	CTG	CGA	CAG	TTT	2454

S	G	S	L	K	P	E	D	A	E	V	F	K	S	P	A	A	S	G	E	838
CA	GGG	TCT	CTG	AAG	CCA	GAG	GAC	GCT	GAG	GTC	TTC	AAG	AGT	CCT	GCC	GCT	TCC	GGG	GAG	2514

K	*																			840
AG	TGA																			2520

FIGURE 1 (cont'd)

GGACGTCACGCAGACAGCACTGTCAACACTGGGCCTTAGGAGACCCCGTTGCCACGGGGGGCTGCTGAGGGAACACCAG

TGCTCTGTCAGCAGCCTGGCCTGGTCTGTGCCTGCCCAGCATGTTCCCAAATCTGTGCTGGACAAGCTGTGGGAAGCGT

TCTTGGAAGCATGGGGAGTGATGTACATCCAACCGTCACTGTCCCCAAGTGAATCTCCTAACAGACTTTCAGGTTTTTA

CTCACTTTACTAAAAAAAAAAAAAAAAAAAGGGCGGCCGCTTA

FIGURE 1 (cont'd)

Full-length human VR2

Input file Flh21e11.seq; Output File Flh21e11.tra
Sequence length 2809

```
GGCTAGCCTGTCCTGACAGGGGAGAGTTAAGCTCCCGTTCTCCACCGTGCCGGCTGGCCAGGTGGGCTGAGGGTGACCG

AGAGACCAGAACCTGCTTGCTGGAGCTTAGTGCTCAGAGCTGGGGAGGGAGGTTCCGCCGCTCCTCTGCTGTCAGCGCC

GGCAGCCCCTCCCGGCTTCACTTCCTCCCGCAGCCCCTGCTACTGAGAAGCTCCGGGATCCCAGCAGCCGCCACGCCCT

GGCCTCAGCCTGCGGGGCTCCAGTCAGGCCAACACCGACGCGCAGCTGGGAGGAAGACAGGACCCTTGACATCTCCATC
```

							M	T	S	P	S	S	S	P		8
TGCACAGAGGTCCTGGCTGGACCGAGCAGCCTCCTCCTCCTAGG							ATG	ACC	TCA	CCC	TCC	AGC	TCT	CCA		24

V	F	R	L	E	T	L	D	G	G	Q	E	D	G	S	E	A	D	R	G	28
GTT	TTC	AGG	TTG	GAG	ACA	TTA	GAT	GGA	GGC	CAA	GAA	GAT	GGC	TCT	GAG	GCG	GAC	AGA	GGA	84

K	L	D	F	G	S	G	L	P	P	M	E	S	Q	F	Q	G	E	D	R	48
AAG	CTG	GAT	TTT	GGG	AGC	GGG	CTG	CCT	CCC	ATG	GAG	TCA	CAG	TTC	CAG	GGC	GAG	GAC	CGG	144

K	F	A	P	Q	I	R	V	N	L	N	Y	R	K	G	T	G	A	S	Q	68
AAA	TTC	GCC	CCT	CAG	ATA	AGA	GTC	AAC	CTC	AAC	TAC	CGA	AAG	GGA	ACA	GGT	GCC	AGT	CAG	204

P	D	P	N	R	F	D	R	D	R	L	F	N	A	V	S	R	G	V	P	88
CCG	GAT	CCA	AAC	CGA	TTT	GAC	CGA	GAT	CGG	CTC	TTC	AAT	GCG	GTC	TCC	CGG	GGT	GTC	CCC	264

E	D	L	A	G	L	P	E	Y	L	S	K	T	S	K	Y	L	T	D	S	108
GAG	GAT	CTG	GCT	GGA	CTT	CCA	GAG	TAC	CTG	AGC	AAG	ACC	AGC	AAG	TAC	CTC	ACC	GAC	TCG	324

E	Y	T	E	G	S	T	G	K	T	C	L	M	K	A	V	L	N	L	K	128
GAA	TAC	ACA	GAG	GGC	TCC	ACA	GGT	AAG	ACG	TGC	CTG	ATG	AAG	GCT	GTG	CTG	AAC	CTT	AAG	384

D	G	V	N	A	C	I	L	P	L	L	Q	I	D	R	D	S	G	N	P	148
GAC	GGA	GTC	AAT	GCC	TGC	ATT	CTG	CCA	CTG	CTG	CAG	ATC	GAC	AGG	GAC	TCT	GGC	AAT	CCT	444

Q	P	L	V	N	A	Q	C	T	D	D	Y	Y	R	G	H	S	A	L	H	168
CAG	CCC	CTG	GTA	AAT	GCC	CAG	TGC	ACA	GAT	GAC	TAT	TAC	CGA	GGC	CAC	AGC	GCT	CTG	CAC	504

I	A	I	E	K	R	S	L	Q	C	V	K	L	L	V	E	N	G	A	N	188
ATC	GCC	ATT	GAG	AAG	AGG	AGT	CTG	CAG	TGT	GTG	AAG	CTC	CTG	GTG	GAG	AAT	GGG	GCC	AAT	564

V	H	A	R	A	C	G	R	F	F	Q	K	G	Q	G	T	C	F	Y	F	208
GTG	CAT	GCC	CGG	GCC	TGC	GGC	CGC	TTC	TTC	CAG	AAG	GGC	CAA	GGG	ACT	TGC	TTT	TAT	TTC	624

G	E	L	P	L	S	L	A	A	C	T	K	Q	W	D	V	V	S	Y	L	228
GGT	GAG	CTA	CCC	CTC	TCT	TTG	GCC	GCT	TGC	ACC	AAG	CAG	TGG	GAT	GTG	GTA	AGC	TAC	CTC	684

L	E	N	P	H	Q	P	A	S	L	Q	A	T	D	S	Q	G	N	T	V	248
CTG	GAG	AAC	CCA	CAC	CAG	CCC	GCC	AGC	CTG	CAG	GCC	ACT	GAC	TCC	CAG	GGC	AAC	ACA	GTC	744

L	H	A	L	V	M	I	S	D	N	S	A	E	N	I	A	L	V	T	S	268
CTG	CAT	GCC	CTA	GTG	ATG	ATC	TCG	GAC	AAC	TCA	GCT	GAG	AAC	ATT	GCA	CTG	GTG	ACC	AGC	804

M	Y	D	G	L	L	Q	A	G	A	R	L	C	P	T	V	Q	L	E	D	288
ATG	TAT	GAT	GGG	CTC	CTC	CAA	GCT	GGG	GCC	CGC	CTC	TGC	CCT	ACC	GTG	CAG	CTT	GAG	GAC	864

I	R	N	L	Q	D	L	T	P	L	K	L	A	A	K	E	G	K	I	E	308

FIGURE 2

ATC CGC AAC CTG CAG GAT CTC ACG ● CTG AAG CTG GCC GCC AAG GAG G●G ATC GAG   924

```
  I   F   R   H   I   L   Q   R   E   F   S   G   L   S   H   L   S   R   K   F    328
ATT TTC AGG CAC ATC CTG CAG CGG GAG TTT TCA GGA CTG AGC CAC CTT TCC CGA AAG TTC   984

  T   E   W   C   Y   G   P   V   R   V   S   L   Y   D   L   A   S   V   D   S    348
ACC GAG TGG TGC TAT GGG CCT GTC CGG GTG TCG CTG TAT GAC CTG GCT TCT GTG GAC AGC  1044

  C   E   E   N   S   V   L   E   I   I   A   F   H   C   K   S   P   H   R   H    368
TGT GAG GAG AAC TCA GTG CTG GAG ATC ATT GCC TTT CAT TGC AAG AGC CCG CAC CGA CAC  1104

  R   M   V   V   L   E   P   L   N   K   L   L   Q   A   K   W   D   L   L   I    388
CGA ATG GTC GTT TTG GAG CCC CTG AAC AAA CTG CTG CAG GCG AAA TGG GAT CTG CTC ATC  1164

  P   K   F   F   L   N   F   L   C   N   L   I   Y   M   F   I   F   T   A   V    408
CCC AAG TTC TTC TTA AAC TTC CTG TGT AAT CTG ATC TAC ATG TTC ATC TTC ACC GCT GTT  1224

  A   Y   H   Q   P   T   L   K   K   Q   A   A   P   H   L   K   A   E   V   G    428
GCC TAC CAT CAG CCT ACC CTG AAG AAG CAG GCC GCC CCT CAC CTG AAA GCG GAG GTT GGA  1284

  N   S   M   L   L   T   G   H   I   L   I   L   L   G   G   I   Y   L   L   V    448
AAC TCC ATG CTG CTG ACG GGC CAC ATC CTT ATC CTG CTA GGG GGG ATC TAC CTC CTC GTG  1344

  G   Q   L   W   Y   F   W   R   R   H   V   F   I   W   I   S   F   I   D   S    468
GGC CAG CTG TGG TAC TTC TGG CGG CGC CAC GTG TTC ATC TGG ATC TCG TTC ATA GAC AGC  1404

  Y   F   E   I   L   F   L   F   Q   A   L   L   T   V   V   S   Q   V   L   C    488
TAC TTT GAA ATC CTC TTC CTG TTC CAG GCC CTG CTC ACA GTG GTG TCC CAG GTG CTG TGT  1464

  F   L   A   I   E   W   Y   L   P   L   L   V   S   A   L   V   L   G   W   L    508
TTC CTG GCC ATC GAG TGG TAC CTG CCC CTG CTT GTG TCT GCG CTG GTG CTG GGC TGG CTG  1524

  N   L   L   Y   Y   T   R   G   F   Q   H   T   G   I   Y   S   V   M   I   Q    528
AAC CTG CTT TAC TAT ACA CGT GGC TTC CAG CAC ACA GGC ATC TAC AGT GTC ATG ATC CAG  1584

  K   V   I   L   R   D   L   L   R   F   L   L   I   Y   L   V   F   L   F   G    548
AAG GTC ATC CTG CGG GAC CTG CTG CGC TTC CTT CTG ATC TAC TTA GTC TTC CTT TTC GGC  1644

  F   A   V   A   L   V   S   L   S   Q   E   A   W   R   P   E   A   P   T   G    568
TTC GCT GTA GCC CTG GTG AGC CTG AGC CAG GAG GCT TGG CGC CCC GAA GCT CCT ACA GGC  1704

  P   N   A   T   E   S   V   Q   P   M   E   G   Q   E   D   E   G   N   G   A    588
CCC AAT GCC ACA GAG TCA GTG CAG CCC ATG GAG GGA CAG GAG GAC GAG GGC AAC GGG GCC  1764

  Q   Y   R   G   I   L   E   A   S   L   E   L   F   K   F   T   I   G   M   G    608
CAG TAC AGG GGT ATC CTG GAA GCC TCC TTG GAG CTC TTC AAA TTC ACC ATC GGC ATG GGC  1824

  E   L   A   F   Q   E   Q   L   H   F   R   G   M   V   L   L   L   L   L   A    628
GAG CTG GCC TTC CAG GAG CAG CTG CAC TTC CGC GGC ATG GTG CTG CTG CTG CTG CTG GCC  1884

  Y   V   L   L   T   Y   I   L   L   L   N   M   L   I   A   L   M   S   E   T    648
TAC GTG CTG CTC ACC TAC ATC CTG CTG CTC AAC ATG CTC ATC GCC CTC ATG AGC GAG ACC  1944

  V   N   S   V   A   T   D   S   W   S   I   W   K   L   Q   K   A   I   S   V    668
GTC AAC AGT GTC GCC ACT GAC AGC TGG AGC ATC TGG AAG CTG CAG AAA GCC ATC TCT GTC  2004

  L   E   M   E   N   G   Y   W   W   C   R   K   K   Q   R   A   G   V   M   L    688
CTG GAG ATG GAG AAT GGC TAT TGG TGG TGC AGG AAG AAG CAG CGG GCA GGT GTG ATG CTG  2064

  T   V   G   T   K   P   D   G   S   P   D   E   R   W   C   F   R   V   E   E    708
```

FIGURE 2 (continued)

ACC GTT GGC ACT AAG CCA GAT GG● ●C CCG GAT GAG CGC TGG TGC TTC A● ●G GAG GAG 2124

 V   N   W   A   S   W   E   Q   T   L   P   T   L   C   E   D   P   S   G   A   728
GTG AAC TGG GCT TCA TGG GAG CAG ACG CTG CCT ACG CTG TGT GAG GAC CCG TCA GGG GCA 2184

 G   V   P   R   T   L   E   N   P   V   L   A   S   P   P   K   E   D   E   D   748
GGT GTC CCT CGA ACT CTC GAG AAC CCT GTC CTG GCT TCC CCT CCC AAG GAG GAT GAG GAT 2244

 G   A   S   E   E   N   Y   V   P   V   Q   L   L   Q   S   N   *   765
GGT GCC TCT GAG GAA AAC TAT GTG CCC GTC CAG CTC CTC CAG TCC AAC TGA   2295

TGGCCCAGATGCAGCAGGAGGCCAGAGGACAGAGCAGAGGATCTTTCCAACCACATCTGCTGGCTCTGGGGTCCCAGTG

AATTCTGGTGGCAAATATATATTTTCACTAACTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 2 (continued)

Partial human VR2 alternate form

Input file frhob12c4.seq; Output File frh b12c4.tra
Sequence length 1489

```
        G   R   F   F   Q   K   G   Q   G   T   C   F   Y   F   G   E   L   P   L        19
GC GGC CGC TTC TTC CAG AAG GGC CAA GGG ACT TGC TTT TAT TTC GGT GAG CTA CCC CTC         57

    S   L   A   A   C   T   K   Q   W   D   V   V   S   Y   L   L   E   N   P   H        39
TCT TTG GCC GCT TGC ACC AAG CAG TGG GAT GTG GTA AGC TAC CTC CTG GAG AAC CCA CAC        117

    Q   P   A   S   L   Q   A   T   D   S   Q   G   N   T   V   L   H   A   L   V        59
CAG CCC GCC AGC CTG CAG GCC ACT GAC TCC CAG GGC AAC ACA GTC CTG CAT GCC CTA GTG        177

    M   I   S   D   N   S   A   E   N   I   A   L   V   T   S   M   Y   D   G   L        79
ATG ATC TCG GAC AAC TCA GCT GAG AAC ATT GCA CTG GTG ACC AGC ATG TAT GAT GGG CTC        237

    L   Q   A   G   A   R   L   C   P   T   V   Q   L   E   D   I   R   N   L   Q        99
CTC CAA GCT GGG GCC CGC CTC TGC CCT ACC GTG CAG CTT GAG GAC ATC CGC AAC CTG CAG        297

    D   L   T   P   L   K   L   A   A   K   E   G   K   I   E   I   F   R   H   I       119
GAT CTC ACG CCT CTG AAG CTG GCC GCC AAG GAG GGC AAG ATC GAG ATT TTC AGG CAC ATC        357

    L   Q   R   E   F   S   G   L   S   H   L   S   R   K   F   T   E   W   C   Y       139
CTG CAG CGG GAG TTT TCA GGA CTG AGC CAC CTT TCC CGA AAG TTC ACC GAG TGG TGC TAT        417

    G   P   V   R   V   S   L   Y   D   L   A   S   V   D   S   C   E   E   N   S       159
GGG CCT GTC CGG GTG TCG CTG TAT GAC CTG GCT TCT GTG GAC AGC TGT GAG GAG AAC TCA        477

    V   L   E   I   I   A   F   H   C   K   S   P   H   R   H   R   M   V   V   L       179
GTG CTG GAG ATC ATT GCC TTT CAT TGC AAG AGC CCG CAC CGA CAC CGA ATG GTC GTT TTG        537

    E   P   L   N   K   L   L   Q   A   K   W   D   L   L   I   P   K   F   F   L       199
GAG CCC CTG AAC AAA CTG CTG CAG GCG AAA TGG GAT CTG CTC ATC CCC AAG TTC TTC TTA        597

    N   F   L   C   N   L   I   Y   M   F   I   F   T   A   V   A   Y   H   Q   P       219
AAC TTC CTG TGT AAT CTG ATC TAC ATG TTC ATC TTC ACC GCT GTT GCC TAC CAT CAG CCT        657

    T   L   K   K   Q   A   A   P   H   L   K   A   E   V   G   N   S   M   L   L       239
ACC CTG AAG AAG CAG GCC GCC CCT CAC CTG AAA GCG GAG GTT GGA AAC TCC ATG CTG CTG        717

    T   G   H   I   L   I   L   L   G   G   I   Y   L   L   V   G   Q   L   W   Y       259
ACG GGC CAC ATC CTT ATC CTG CTA GGG GGG ATC TAC CTC CTC GTG GGC CAG CTG TGG TAC        777

    F   W   R   R   H   V   F   I   W   I   S   F   I   D   S   Y   F   E   I   L       279
TTC TGG CGG CGC CAC GTG TTC ATC TGG ATC TCG TTC ATA GAC AGC TAC TTT GAA ATC CTC        837

    F   L   F   Q   A   L   L   T   V   V   S   Q   V   L   C   F   L   A   I   E       299
TTC CTG TTC CAG GCC CTG CTC ACA GTG GTG TCC CAG GTG CTG TGT TTC CTG GCC ATC GAG        897

    W   Y   L   P   L   L   V   S   A   L   V   L   G   W   L   N   L   L   Y   Y       319
TGG TAC CTG CCC CTG CTT GTG TCT GCG CTG GTG CTG GGC TGG CTG AAC CTG CTT TAC TAT        957

    T   R   G   F   Q   H   T   G   I   Y   S   V   M   I   Q   K   K   A   I   S       339
ACA CGT GGC TTC CAG CAC ACA GGC ATC TAC AGT GTC ATG ATC CAG AAG AAA GCC ATC TCT       1017

    V   L   E   M   E   N   G   Y   W   W   C   R   K   K   Q   R   A   G   V   M       359
GTC CTG GAG ATG GAG AAT GGC TAT TGG TGG TGC AGG AAG AAG CAG CGG GCA GGT GTG ATG       1077
```

FIGURE 3

```
  L   T   V   G   T   K   P   D ● G   S   P   D   E   R   W   C ●  R   V   E   379
CTG ACC GTT GGC ACT AAG CCA GAT GGC AGC CCG GAT GAG CGC TGG TGC TTC AGG GTG GAG  1137

  E   V   N   W   A   S   W   E   Q   T   L   P   T   L   C   E   D   P   S   G   399
GAG GTG AAC TGG GCT TCA TGG GAG CAG ACG CTG CCT ACG CTG TGT GAG GAC CCG TCA GGG  1197

  A   G   V   P   R   T   L   E   N   P   V   L   A   S   P   P   K   E   D   E   419
GCA GGT GTC CCT CGA ACT CTC GAG AAC CCT GTC CTG GCT TCC CCT CCC AAG GAG GAT GAG  1257

  D   G   A   S   E   E   N   Y   V   P   V   Q   L   L   Q   S   N   *          437
GAT GGT GCC TCT GAG GAA AAC TAT GTG CCC GTC CAG CTC CTC CAG TCC AAC TGA          1311
```

TGGCCCAGATGCAGCAGGAGGCCAGAGGACAGAGCAGAGGATCTTTCCAACCACATCTGCTGGCTCTGGGGTCCCAGTG

AATTCTGGTGGCAAATATATATTTTCACTAACTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGA

GCGGACGCGTGGGTCGAC

Partial rat VR2

Input file Flrxb147g11.seq; Output File Flrxb147g11.tra
Sequence length 1794

```
      S    T    H    A    S    A    L    S    L    A    A    C    T    K    Q    W    D    V    V      19
G TCG ACC CAC GCG TCC GCT CTT TCT CTG GCT GCG TGC ACC AAG CAG TGG GAT GTG GTG      57

      T    Y    L    L    E    N    P    H    Q    P    A    S    L    E    A    T    D    S    L    G   39
ACC TAC CTC CTG GAG AAC CCA CAC CAG CCG GCC AGC CTG GAG GCC ACC GAC TCC CTG GGC    117

      N    T    V    L    H    A    L    V    M    I    A    D    N    S    P    E    N    S    A    L   59
AAC ACA GTC CTG CAT GCT CTG GTA ATG ATT GCA GAT AAC TCG CCT GAG AAC AGT GCC CTG    177

      V    I    H    M    Y    D    G    L    L    Q    M    G    A    R    L    C    P    T    V    Q   79
GTG ATC CAC ATG TAC GAC GGG CTT CTA CAA ATG GGG GCG CGC CTC TGC CCC ACT GTG CAG    237

      L    E    E    I    S    N    H    Q    G    L    T    P    L    K    L    A    A    K    E    G   99
CTT GAG GAA ATC TCC AAC CAC CAA GGC CTC ACA CCC CTG AAA CTA GCC GCC AAG GAA GGC    297

      K    I    E    I    F    R    H    I    L    Q    R    E    F    S    G    P    Y    Q    P    L  119
AAA ATC GAG ATT TTC AGG CAC ATT CTG CAG CGG GAA TTC TCA GGA CCG TAC CAG CCC CTT    357

      S    R    K    F    T    E    W    C    Y    G    P    V    R    V    S    L    Y    D    L    S  139
TCC CGA AAG TTT ACT GAG TGG TGT TAC GGT CCT GTG CGG GTA TCG CTG TAC GAC CTG TCC    417

      S    V    D    S    W    E    K    N    S    V    L    E    I    I    A    F    H    C    K    S  159
TCT GTG GAC AGC TGG GAA AAG AAC TCG GTG CTG GAG ATC ATC GCT TTT CAT TGC AAG AGC    477

      P    N    R    H    R    M    V    V    L    E    P    L    N    K    L    L    Q    E    K    W  179
CCG AAC CGG CAC CGC ATG GTG GTT TTA GAA CCA CTG AAC AAG CTT CTG CAG GAG AAA TGG    537

      D    R    L    V    S    R    F    F    F    N    F    A    C    Y    L    V    Y    M    F    I  199
GAT CGG CTC GTC TCA AGA TTC TTC TTC AAC TTC GCC TGC TAC TTG GTC TAC ATG TTC ATC    597

      F    T    V    V    A    Y    H    Q    P    S    L    D    Q    P    A    I    P    S    S    K  219
TTC ACC GTC GTT GCC TAC CAC CAG CCT TCC CTG GAT CAG CCA GCC ATC CCC TCA TCA AAA    657

      A    T    F    G    E    S    M    L    L    L    G    H    I    L    I    L    L    G    G    I  239
GCG ACT TTT GGG GAA TCC ATG CTG CTG CTG GGC CAC ATT CTG ATC CTG CTT GGG GGT ATT    717

      Y    L    L    L    G    Q    L    W    Y    F    W    R    R    R    L    F    I    W    I    S  259
TAC CTC TTA CTG GGC CAG CTG TGG TAC TTT TGG CGG CGG CGC CTG TTT ATC TGG ATC TCA    777

      F    M    D    S    Y    F    E    I    L    F    L    L    Q    A    L    L    T    V    L    S  279
TTC ATG GAC AGC TAC TTT GAA ATC CTC TTT CTC CTT CAG GCT CTG CTC ACA GTG CTG TCC    837

      Q    V    L    R    F    M    E    T    E    W    Y    L    P    L    L    V    L    S    L    V  299
CAG GTG CTG CGC TTC ATG GAG ACT GAA TGG TAC CTA CCC CTG CTA GTG TTA TCC CTA GTG    897

      L    G    W    L    N    L    L    Y    Y    T    R    G    F    Q    H    T    G    I    Y    S  319
CTG GGC TGG CTG AAC CTG CTT TAC TAC ACA CGG GGC TTT CAG CAC ACA GGC ATC TAC AGT    957

      V    M    I    Q    K    V    I    L    R    D    L    L    R    F    L    L    V    Y    L    V  339
GTC ATG ATC CAG AAG GTC ATC CTT CGA GAC CTG CTC CGT TTC CTG CTG GTC TAC CTG GTC   1017

      F    L    F    G    F    A    V    A    L    V    S    L    S    R    E    A    R    S    P    K  359
TTC CTT TTC GGC TTT GCT GTA GCC CTA GTA AGC TTG AGC AGA GAG GCC CGA AGT CCC AAA   1077
```

FIGURE 4

```
  A   P   E   D   N   N   S   T   ●   T   E   Q   P   T   V   G   ●   E   E   E    379
GCC CCT GAA GAT AAC AAC TCC ACA GTG ACG GAA CAG CCC ACG GTG GGC CAG GAG GAG GAG  1137

  P   A   P   Y   R   S   I   L   D   A   S   L   E   L   F   K   F   T   I   G    399
CCA GCT CCA TAT CGG AGC ATT CTG GAT GCC TCC CTA GAG CTG TTC AAG TTC ACC ATT GGT  1197

  M   G   E   L   A   F   Q   E   Q   L   R   F   R   G   V   V   L   L   L   L    419
ATG GGG GAG CTG GCT TTC CAG GAA CAG CTG CGT TTT CGT GGG GTG GTC CTG CTG TTG CTG  1257

  L   A   Y   V   L   L   T   Y   V   L   L   L   N   M   L   I   A   L   M   S    439
TTG GCC TAC GTC CTT CTC ACC TAC GTC CTG CTG CTC AAC ATG CTC ATT GCT CTC ATG AGC  1317

  E   T   V   N   H   V   A   D   N   S   W   S   I   W   K   L   Q   K   A   I    459
GAA ACT GTC AAC CAC GTT GCT GAC AAC AGC TGG AGC ATC TGG AAG TTG CAG AAA GCC ATC  1377

  S   V   L   E   M   E   N   G   Y   W   W   C   R   R   K   K   H   R   E   G    479
TCT GTC TTG GAG ATG GAG AAT GGT TAC TGG TGG TGC CGG AGG AAG AAA CAT CGT GAA GGG  1437

  R   L   L   K   V   G   T   R   G   D   G   T   P   D   E   R   W   C   F   R    499
AGG CTG CTG AAA GTC GGC ACC AGG GGG GAT GGT ACC CCT GAT GAG CGC TGG TGC TTC AGG  1497

  V   E   E   V   N   W   A   A   W   E   K   T   L   P   T   L   S   E   D   P    519
GTG GAG GAA GTA AAT TGG GCT GCT TGG GAG AAG ACT CTT CCC ACC TTA TCT GAG GAT CCA  1557

  S   G   P   G   I   T   G   N   K   K   N   P   T   S   K   P   G   K   N   S    539
TCA GGG CCA GGC ATC ACT GGT AAT AAA AAG AAC CCA ACC TCT AAA CCG GGG AAG AAC AGT  1617

  A   S   E   E   D   H   L   P   L   Q   V   L   Q   S   P   *                    555
GCC TCA GAG GAA GAC CAT CTG CCC CTT CAG GTC CTC CAG TCC CCC TGA                   1665

TGGCCCAGATGCAGCAGCAGGCTGGCAGGATGGAGTAGGGAATCTTCCCAGCCACACCAGAGGCTACTGAATTTTGGTG

GAAATATAAATATTTTTTTTTGCATAAAAAAAAAAAAAAAAAAAGGGCGGCCGC
```

FIGURE 4 (continued)

GAP of: humanvr2.pep   check: 5746   from: 1   to: 764

humanVR2 Flh21e11

 to: humanvr1.pep   check: 6877   from: 1   to: 839

humanVR1 _Fbh18547pat - fchrb87a6, 3909 bases, 4554 checksum.

 Symbol comparison table:
/ddm_local/gcg/gcg_9.1/gcgcore/data/rundata/blosum62.cmp
 CompCheck: 6430

            Gap Weight: 12   Average Match: 2.912
         Length Weight: 4   Average Mismatch: -2.003

              Quality: 1530   Length: 850
                Ratio: 2.003   Gaps: 10
   Percent Similarity: 55.378   Percent Identity: 46.348

        Match display thresholds for the alignment(s):
                   | = IDENTITY
                   : =   2
                   . =   1

humanvr2.pep x humanvr1.pep

```
   1 ...............................................MTSPSSSPVF  10
                                                   |. |  ..|
   1 MKKWSSTDLGTAADPLQKDTCPDPLDGDPNSRPPPAKPQLPTAKSRTRLF  50

  11 RLETLDGGQEDGSEADRGKLDFGSGLPPMESQFQGEDRKFAPQIRVNLNY  60
       :     : |.||    :    |   .       |.
  51 GKGDSEEAFPVDCPHEEGELDSCPTI.TVSPVITIQRPGDGPTGARLLSQ  99

  61 RKGTGASQPDPNRFDRDRLFNAVSRGVPEDLAGLPEYLSKTSKYLTDSEY 110
      ..:         :||  :| ||..   :||   |  :| |. |:|||.|:
 100 DSVAASTEKTLRLYDRRSIFEAVAQNNCQDLESLLLFLQKSKKHLTDNEF 149

 111 TEGSTGKTCLMKAVLNLKDGVNACILPLLQIDRDSGNPQPLVNAQCTDDY 160
       :  ||||||:||.||| || |  |   ||:|  . . |||| || |
 150 KDPETGKTCLLKAMLNLHDGQNTTIPLLLEIARQTDSLKELVNASYTDSY 199

 161 YRGHSALHIAIEKRSLQCVKLLVENGANVHARACGRFFQKGQG.TCFYFG 209
     |:|  .||||||||:|.:  |  ||||||||.| | | | ||.| .|   ||||
 200 YKGQTALHIAIERRNMALVTLLVENGADVQAAAHGDFFKKTKGRPGFYFG 249

 210 ELPLSLAACTKQWDVVSYLLENPHQPASLQATDSQGNTVLHALVMISDNS 259
     ||||||||||| |  :| :||:|  | | : | || ||.|||||| :.||.
 250 ELPLSLAACTNQLGIVKFLLQNSWQTADISARDSVGNTVLHALVEVADNT 299

 260 AENIALVTSMYDGLLQAGARLCPTVQLEDIRNLQDLTPLKLAAKEGKIEI 309
     |:|   |||||. :|  ||:| ||..||:: |  . :|||.||| |||  :
 300 ADNTKFVTSMYNEILMLGAKLHPTLKLEELTNKKGMTPLALAAGTGKIGV 349
```

<div align="center">

FIGURE 5

</div>

```
310 FRHILQREFS..GLSHLSRKFTEWCYGPVRVSLYDLASVDSCEENSVLEI 357
    :|||||     |||||||||| |||| |||||. :|.||.|||||:
350 LAYILQREIQEPECRHLSRKFTEWAYGPVHSSLYDLSCIDTCEKNSVLEV 399

358 IAF.HCKSPHRHRMVVLEPLNKLLQAKWDLLIPK.FFLNFLCNLIYMFIF 405
    ||:   ..|.|| |...||||:||| |||  : : |: |||   :|| ||
400 IAYSSSETPNRHDMLLVEPLNRLLQDKWDRFVKRIFYFNFLVYCLYMIIF 449

406 TAVAYHQPTLKKQAAPHLKAEVGNSMLLTGHILILLGGIYLLVGQLWYFW 455
    |  ||:.|    ..|      .:|.  .|| || .|||:|     : ||
450 TMAAYYRPV..DGLPPFKMEKIGDYFRVTGEILSVLGGVYFFFRGIQYFL 497

456 RRHVFIWISFIDSYFEILFLFQALLTVVSQVLCFLAIEWYLPLLVSALVL 505
    .|    .   |:||| |.|| |.|  . . || | :. |. :| .| |
498 QRRPSMKTLFVDSYSEMLFFLQSLFMLATVVLYFSHLKEYVASMVFSLAL 547

506 GWLNLLYYTRGFQHTGIYSVMIQKVILRDLLRFLLIYLVFLFGFAVALVS 555
    || |:||||||||  |||.|||:|.|||||| ||: :|:||||||. |.|.
548 GWTNMLYYTRGFQQMGIYAVMIEKMILRDLCRFMFVYIVFLFGFSTAVVT 597

556 LSQEAWRPEAPTGPNATESVQPMEGQEDEGNGAQYRGILEASLELFKFTL-605
    | ::      |.   .     |     |     . |  :     |||||||||
598 LIEDGKNDSLPSESTSHRWRGPACRPPD....SSYNSLYSTCLELFKFTI 643

606 GMGELAFQEQLHFRGMVLLLLLAYVLLTYILLLNMLIALMSETVNSVATD 655
    |||:| | |    |: . ::|||||||:|||||||||||||| |||| :| :
644 GMGDLEFTENYDFKAVFIILLLAYVILTYILLLNMLIALMGETVNKIAQE 693

656 SWSIWKLQKAISVLEMENGYWWC.RKKQRAGVMLTVGTKPDGSPDERWCF 704
    | .|||||:||.:|: | : | || |.| :| || ||| | ||||
694 SKNIWKLQRAITILDTEKSFLKCMRKAFRSGKLLQVGYTPDGKDDYRWCF 743

705 RVEEVNWASWEQTLPTLCEDPSGA.GVPRTLENPVLASPPKEDEDGASEE 753
    ||:||||  .|    .   : ||| || |||  . .|         |  .
744 RVDEVNWTTWNTNVGIINEDPGNCEGVKRTLSFSLRSS....RVSGRHWK 789

754 NYVPVQLLQSN..................................... 764
    |:  | ||.
790 NFALVPLLREASARDRQSAQPEEVYLRQFSGSLKPEDAEVFKSPAASGEK 839
```

FIGURE 5 (cont'd)

GAP of: humanvr2.seq   check: 8853   from: 1   to: 2809

humanVR2 21e11a, 2809 bases, 8853 checksum.

 to: humanvr1.seq   check: 4554   from: 1   to: 3909

humanVR1 Fbh18547pat - Import - complete

 Symbol comparison table:
/ddm_local/gcg/gcg_9.1/gcgcore/data/rundata/nwsgapdna.cmp
 CompCheck: 8760

          Gap Weight: 50   Average Match: 10.000
          Length Weight: 3   Average Mismatch: 0.000

          Quality: 14359   Length: 3934
          Ratio: 5.112   Gaps: 15
Percent Similarity: 55.316   Percent Identity: 55.316

          Match display thresholds for the alignment(s):
                    | = IDENTITY
                    : =   5
                    . =   1

humanvr2.seq x humanvr1.seq

     1 .......................GGCTAGCCTGTCCTGACAGGGGAGAG 26
                             | | | | | || |||
   801 TGTCCACAGTAGTCCCCCCTTATCCACGGGTGTCACTTTCCATGGGTTCA 850

    27 TTAAGCTCCCGTTCTCCACCGTGCCGGCTGGCCAGGTGGGCTGAGGGTGA 76
       | | | | || | |||| || | ||| | | |
   851 GTTATTTGCGGTCAACCACGGTCTGCCAATATTAAATGGAAAATTCTTCA 900

    77 CCGAGAGACCAGAACCTGCTTGCTGGAGCTTAGTGCTCAGAGCTGGGGAG 126
          || ||| | | | || | | || ||| | | ||
   901 AACAGTTCCCAAGTTTTCCCTTGTGCATTGTTCTGAGCAGTGTGATGAAG 950

   127 GGAGGTTCCGCCGCTCCTCTGCTGTCAGCGCCGGCAGCCCCTCCCGGCTT 176
       | | | | || | .|| | | ||| | | |
   951 AGTCTCTGCCGTGCCATCTGGGATGCAAACCGTCCCTGTGTCCCCCACGT 1000

   177 CACTTCCTCCCGCAGCCCCTGCTACTGAGAAGCTCCGGGATCCCAGCAGC 226
       | || ||| || | || || | || || | | |
  1001 CCAGGCCGTAGATGCTCCCCGCCGGTCAGTCACTTAGTCGTCAGATCGCC 1050

   227 CGCCACGCCCTGGC.....................CTCAGCCTGCGGG 253
       . || | | | | || ||| | ||
  1051 CGTCCTGGTATCACAGTGCTTCTGTTCAGGTTGCACACTGGGCCACAGAG 1100

FIGURE 6

```
 254 GCTCCAGTCAGGCCAACACCGACGCGCAGCTGGGAGGAAG.......... 293
     I IIIII  III   I      I IIIII  II         I
1101 GATCCAGCAAGGATGAAGAAATGGAGCAGCACAGACTTGGGGACAGCTGC 1150

 294 ...........ACAGGACCCTTGACATCTCCATCTGCACAGAGGTCCTG 331
                I IIIII I II I    II  III I  III  III
1151 GGACCCACTCCAAAAGGACACCTGCCCAGACCCCCTGGATGGAGACCCTA 1200

 332 GCTGGACCGAGCAGCCTCCTCCTCCTAGGATGACCTCACCCTCCAGC..T 379
      II I    I  II I    II    I I  II I   II    III
1201 ACTCCAGGCCACCTCCAGCCAAGCCCCAGCTCCCCACGGCCAAGAGCCGC 1250

 380 CTCCAGTTTTCAGGTTGGAGACATTAGATGGAGGCCAAGAAGATGGCTCT 429
      II I I  II   I I  I I I  IIIIII      I III I
1251 ACCCGGCTCTTTGGGAAGGGTGACTCGGAGGAGGCTTTCCCGGTGGATTG 1300

 430 GAGGCGGACAGAGGAAAGCTGGATTTTGGGAGCGGGCTGCCTCCCATGGA 479
      I     II I    I I  I  I  I II I I       I
1301 CCCCCACGAGGAAGGTGAGTTGGACTCCTGCCCGACCATCACAGTCAGCC 1350

 480 GTCACAGTTCCAGGGCGAGGACCGGAAATTCGCCCCTCAGATAAGAGTCA 529
      I  I  III  IIII I III I  III I I I I
1351 CTGTTATCACCATCCAGAGGCCAGGAGACGGCCCCACCGGTGCCAGG..G 1398

 530 ACCTCAACTACCGAAAGGGAACAGGTGCCAGTCAGCCGGATCCAAACCGA 579
      II   I I     I   I I    II   II  I   I   I
1399 TGCTGTCCCAGGACTCTGTCGCCGCCAGCACCGAGAAGACCCTCAGGCTC 1448

 580 TTTGACCGAGATCGGCTCTTCAATGCGGTCTCCCGGGGTGTCCCCGAGGA 629
      I IIII II   I   IIII I II II I I I  I  I I IIII
1449 TATGATCGCAGGAGTATCTTTGAAGCCGTTGCTCAGAATAACTGCCAGGA 1498

 630 TCTGGCTGGACTTCCAGAGTACCTGAGCAAGACCAGCAAGTACCTCACCG 679
      IIIII   I II I    I IIII   IIII II  III IIIIIII I
1499 TCTGGAGAGCCTGCTGCTCTTCCTGCAGAAGAGCAAGAAGCACCTCACAG 1548

 680 ACTCGGAATACACAGAGGGCTCCACAGGTAAGACGTGCCTGATGAAGGCT 729
      II   II I II III      IIIII IIIII II III IIII II
1549 ACAACGAGTTCAAAGACCCTGAGACAGGGAAGACCTGTCTGCTGAAAGCC 1598

 730 GTGCTGAACCTTAAGGACGGAGTCAATGCCTGCATTCTGCCACTGCTGCA 779
      IIII IIIII  I IIIIII   II  II  III I  I  II III I
1599 ATGCTCAACCTGCACGACGGACAGAACACCACCATCCCCCTGCTCCTGGA 1648

 780 GATCGACAGGGACTCTGGCAATCCTCAGCCCCTGGTAAATGCCCAGTGCA 829
      IIIII   II I I I II I   II   II II II III   I II
1649 GATCGCGCGGCAAACGGACAGCCTGAAGGAGCTTGTCAACGCCAGCTACA 1698

 830 CAGATGACTATTACCGAGGCCACAGCGCTCTGCACATCGCCATTGAGAAG 879
      I III   III III    IIIII I   II IIIIIIIIIIIIIII IIII
1699 CGGACAGCTACTACAAGGGCCAGACAGCACTGCACATCGCCATCGAGAGA 1748

 880 AGGAGTCTGCAGTGTGTGAAGCTCCTGGTGGAGAATGGGGCCAATGTGCA 929
      I I   II    IIII IIIIIIIIIIIIIII II II  I II II
1749 CGCAACATGGCCCTGGTGACCCTCCTGGTGGAGAACGGAGCAGACGTCCA 1798

 930 TGCCCGGGCCTGCGGCCGCTTCTTCCAGAAGGGCCAAG...GGACTTGCT 976
      II   IIII  II   IIIIII IIII  I III  II II I I
1799 GGCTGCGGCCCATGGGGACTTCTTTAAGAAAACCAAAGGGCGGCCTGGAT 1848
```

FIGURE 6 (cont'd)

```
 977 TTTATTTCGGTGAGCTACCCCTCTCTTTGGCCGCTTGCACCAAGCAGTGG 1026
      | || ||||||||| || |||||| ||  ||||||| |||||||| ||| |
1849 TCTACTTCGGTGAACTGCCCCTGTCCCTGGCCGCGTGCACCAACCAGCTG 1898

1027 GATGTGGTAAGCTACCTCCTGGAGAACCCACACCAGCCCGCCAGCCTGCA 1076
      |  | ||| | | ||| ||| |||||| |    ||| | |||  | |
1899 GGCATCGTGAAGTTCCTGCTGCAGAACTCCTGGCAGACGGCCGACATCAG 1948

1077 GGCCACTGACTCCCAGGGCAACACAGTCCTGCATGCCCTAGTGATGATCT 1126
      ||||   |||||  .|||||||||| || ||||| ||||| ||| | |
1949 CGCCAGGGACTCGGTGGGCAACACGGTGCTGCACGCCCTGGTGGAGGTGG 1998

1127 CGGACAACTCAGCTGAGAACATTGCACTGGTGACCAGCATGTATGATGGG 1176
      | |||||| | || || ||||       | |||||| ||||||||  ||| |
1999 CCGACAACACGGCCGACAACACGAAGTTTGTGACGAGCATGTACAATGAG 2048

1177 CTCCTCCAAGCTGGGGCCCGCCTCTGCCCTACCGTGCAGCTTGAGGACAT 1226
      | ||        ||||||  ||   ||| ||  || |||| |||||| |
2049 ATTCTGATGCTGGGGGCCAAACTGCACCCGACGCTGAAGCTGGAGGAGCT 2098

1227 CCGCAACCTGCAGGATCTCACGCCTCTGAAGCTGGCCGCCAAGGAGGGCA⁻1276
      |  |||| | |||  |  |||||| |||   |||||| ||  |   || |
2099 CACCAACAAGAAGGGAATGACGCCGCTGGCTCTGGCAGCTGGGACCGGGA 2148

1277 AGATCGAGATTTTCAGGCACATCCTGCAGCGGGAGTT......TTCAGGA 1320
      |||||| | | ||     | || || ||||||||| |      | |
2149 AGATCGGGGTCTTGGCCTATATTCTCCAGCGGGAGATCCAGGAGCCCGAG 2198

1321 CTGAGCCACCTTTCCCGAAAGTTCACCGAGTGGTGCTATGGGCCTGTCCG 1370
      ||  ||||| |||  | |||||||||||||||||||  ||| |||||| || |
2199 TGCAGGCACCTGTCCAGGAAGTTCACCGAGTGGGCCTACGGGCCCGTGCA 2248

1371 GGTGTCGCTGTATGACCTGGCTTCTGTGGACAGCTGTGAGGAGAACTCAG 1420
      |||||||| ||||||| | |    | |||| ||| ||| ||||||||| |
2249 CTCCTCGCTGTACGACCTGTCCTGCATCGACACCTGCGAGAAGAACTCGG 2298

1421 TGCTGGAGATCATTGCCTTTCATTGCA...AGAGCCCGCACCGACACCGA 1467
      ||||||||| | || ||||   |||   ||| ||| | ||| |||
2299 TGCTGGAGGTGATCGCCTACAGCAGCAGCGAGACCCCTAATCGCCACGAC 2348

1468 ATGGTCGTTTTGGAGCCCCTGAACAAACTGCTGCAGGCGAAATGGGA... 1514
      ||| || |  ||||||| ||||||  ||| ||||||||  || |||||
2349 ATGCTCTTGGTGGAGCCGCTGAACCGACTCCTGCAGGACAAGTGGGACAG 2398

1515 TCTGCTCATCCCCAAGTTCTTCTTAAACTTCCTGTGTAATCTGATCTACA 1564
      | |||  | ||  |||| |||  ||||||||| |      | ||||
2399 ATTCGTCAAGCGCATCTTCTACTTCAACTTCCTGGTCTACTGCCTGTACA 2448

1565 TGTTCATCTTCACCGCTGTTGCCTACCATCAGCCTACCCTGAAGAAGCAG 1614
      || |||||||||||   | |||||||| |   |||          ||
2449 TGATCATCTTCACCATGGCTGCCTACTA....CAGGCCCGTGGATGGCTT 2494

1615 GCCGCCCCTCACCTGAAAGCGGAGGTTGGAAACTCCATGCTGCTGACGGG 1664
      ||| ||| | |  || |  |  ||||| ||| | |  |  | || ||
2495 GCCTCCCTTTA..AGATGGAAAAAATTGGAGACTATTTCCGAGTTACTGG 2542
```

FIGURE 6 (cont'd)

```
1665 CCACATCCTTATCCTGCTAGGGGGGATCTACCTCCTCGTGGGCCAGCTGT 1714
     I IIIIII    II IIII II  IIIII II I  I  I    I I
2543 AGAGATCCTGTCTGTGTTAGGAGGAGTCTACTTCTTTTTCCGAGGGATTC 2592

1715 GGTACTTCTGGCGGCGCCACGTGTTCATCTGGATCTCGTTCATAGACAGC 1764
     III III  II I I I    II  II   II I III  I IIIIIII
2593 AGTATTTCCTGCAGAGGCGGCCGTCGATGAAGACCCTGTTTGTGGACAGC 2642

1765 TACTTTGAAATCCTCTTCCTGTTCCAGGCCCTGCTCACAGTGGTGTCCCA 1814
     III  III II II..III I  I III I III III   III  II
2643 TACAGTGAGATGCTTTTCTTTCTGCAGTCACTGTTCATGCTGGCCACCGT 2692

1815 GGTGCTGTGTTTCCTGGCCATCGAGTGGTACCTGCCCCTGCTTGTGTCTG 1864
     IIIIIIII  III    I II II  III  II I    I III I
2693 GGTGCTGTACTTCAGCCACCTCAAGGAGTATGTGGCTTCCATGGTATTCT 2742

1865 CGCTGGTGCTGGGCTGGCTGAACCTGCTTTACTATACACGTGGCTTCCAG 1914
     I IIIII   IIIIIIII   III IIII IIIII II II II IIIIII
2743 CCCTGGCCTTGGGCTGGACCAACATGCTCTACTACACCCGCGGTTTCCAG 2792

1915 CACACAGGCATCTACAGTGTCATGATCCAGAAGGTCATCCTGCGGGACCT 1964
     II I IIIIIIIII    IIIIIIII  IIIII I IIIIII I IIIII
2793 CAGATGGGCATCTATGCCGTCATGATAGAGAAGATGATCCTGAGAGACCT 2842

1965 GCTGCGCTTCCTTCTGATCTACTTAGTCTTCCTTTTCGGCTTCGCTGTAG 2014
     I   II III I  I  IIIII I IIIIII I IIIII II I   II
2843 GTGCCGTTTCATGTTTGTCTACATCGTCTTCTTGTTCGGGTTTTCCACAG 2892

2015 CCCTGGTGAGCCTGAGCCAGGAGGCTTGGCGCCCCGAAGCTCCTACAGGC 2064
     I IIIIIII IIII  I III  I      I      II  I I
2893 CGGTGGTGACGCTGATTGAAGACGGGAAGAATGACTCCCTGCCGTCTGAG 2942

2065 CCCAATGCCACAGAGTCAGTGCAGCCCATGGAGGGACAGGAGGACGAGGG 2114
     III         I   III    I III I I     I I
2943 TCCA..........CGTCGCACAGGTGGCGGGGGCCTGCCTGCAGGCC 2980

2115 CAACGGGGCCCAGTACAGGGGTATCCTGGAAGCCTCCTTGGAGCTCTTCA 2164
     I  II   I   IIII  I  I    III I IIIIIII IIII
2981 CCCCGATAGCTCCTACAACAGCCTGTACTCCACCTGCCTGGAGCTGTTCA 3030

2165 AATTCACCATCGGCATGGGCGAGCTGGCCTTCCAGGAGCAGCTGCACTTC 2214
     I IIIIIIIIIIIIIIIIIIIII IIII  III   III I   IIIII
3031 AGTTCACCATCGGCATGGGCGACCTGGAGTTCACTGAGAACTATGACTTC 3080

2215 CGCGGCATGGTGCTGCTGCTGCTGCTGGCCTACGTGCTGCTCACCTACAT 2264
      I  I  I  I  I IIIIIIIIIIIIII II  I IIIIIIIIIIII
3081 AAGGCTGTCTTCATCATCCTGCTGCTGGCCTATGTAATTCTCACCTACAT 3130

2265 CCTGCTGCTCAACATGCTCATCGCCCTCATGAGCGAGACCGTCAACAGTG 2314
     III IIIIIIIIIIIIIIIIIIIIIIIIIII I IIIIII IIIIIII
3131 CCTCCTGCTCAACATGCTCATCGCCCTCATGGGTGAGACTGTCAACAAGA 3180

2315 TCGCCACTGACAGCTGGAGCATCTGGAAGCTGCAGAAAGCCATCTCTGTC 2364
     IIII    II III  II IIIIIIIIIIIIIIIIII IIIIIII I  II
3181 TCGCACAGGAGAGCAAGAACATCTGGAAGCTGCAGAGAGCCATCACCATC 3230
```

FIGURE 6 (cont'd)

```
2365 CTGGAGATGGAGAATGGCTATTGGTGGTGCAGGAAGAAG...CAGCGGGC 2411
     |||||  | ||||||   |||         ||||| || ||||      || |
3231 CTGGACACGGAGAAGAGCTTCCTTAAGTGCATGAGGAAGGCCTTCCGCTC 3280

2412 AGGTGTGATGCTGACCGTTGGCACTAAGCCAGATGGCAGCCCGGATGAGC 2461
     |||    | ||||||    || ||     |  || |||||||     || | |
3281 AGGCAAGCTGCTGCAGGTGGGGTACACACCTGATGGCAAGGACGACTACC 3330

2462 GCTGGTGCTTCAGGGTGGAGGAGGTGAACTGGGCTTCATGGGAGCAGACG 2511
     | ||||||||||||||||||| |||||||||||| | | ||| |     |
3331 GGTGGTGCTTCAGGGTGGACGAGGTGAACTGGACCACCTGGAACACCAAC 3380

2512 CTGCCTACGCTGTGTGAGGACCCG...TCAGGGCAGGTGTCCCTCGAAC 2558
     ||   |   |    || ||||||      | | || ||| || || ||
3381 GTGGGCATCATCAACGAAGACCCGGGCAACTGTGAGGGCGTCAAGCGCAC 3430

2559 TCTCGAGAACCCTGTCCTG....GCTTCCCCTCCCAAGGAGGATGAGGAT 2604
     ||     | | | | |    ||      ||      ||    ||
3431 CCTGAGCTTCTCCCTGCGGTCAAGCAGAGTTTCAGGCAGACACTGGAAGA 3480

2605 GGTGCCTCTGAGGAAAACTATGTGCCCGTCCAGCTCCTCCAGTCCAACTG 2654
       |     |     |  || |         || | ||| ||.. | |
3481 ACTTTGCCCTGGTCCCCCTTTTAAGAGAGGCAAGTGCTCGAGATAGGCAG 3530

2655 ATGGCCCAGATGCAGCAGGAGGCCAGAGGACAGAGCAGAGGATCTTTCCA 2704
     || |||   || | |        |||||     ||| ||| |  |
3531 TCTGCTCAGCCCGAGGAAGTTTATCTGCGACAGTTTTCAGGGTCTCTGAA 3580

2705 ACCACATCTGCTGGCTCTGGGGTCCCAGTGAATTCTGGTGGCAAATATAT 2754
     |||    |   | |   ||| ||||   || | |||  |
3581 GCCA.....GAGGACGCTGAGGTCTTCAAGAGTCCTGCCGCTTCCGGGGA 3625

2755 ATTTTCACTAACTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2804
     | |  |   ||| | |  |     || |          |   | |
3626 GAAGTGAGGACGTCACGCAGACAGCACTGTCAACACTGGGCCTTAGGAGA 3675

2805 AAAAA.......................................... 2809
3676 CCCCGTTGCCACGGGGGGCTGCTGAGGGAACACCAGTGCTCTGTCAGCAG 3725
```

FIGURE 6 (cont'd)

```
humanVR2    MTSPSSSPVFRLETLDGGQEDGSEADRGKLDFGSGLPPMESQFQGEDRKFAPQIRVNLNY
ratVR2      ------------------------------------------------------------


humanVR2    RKGTGASQPDPNRFDRDRLFNAVSRGVPEDLAGLPEYLSKTSKYLTDSEYTEGSTGKTCL
ratVR2      ------------------------------------------------------------


humanVR2    MKAVLNLKDGVNACILPLLQIDRDSGNPQPLVNAQCTDDYYRGHSALHIAIEKRSLQCVK
ratVR2      ------------------------------------------------------------


humanVR2    LLVENGANVHARACGRFFQKGQGTCFYFGELPLSLAACTKQWDVVSYLLENPHQPASLQA
ratVR2      -----------------------STHASALSLAACTKQWDVVTYLLENPHQPASLEA
                                   .*************:*************:*

humanVR2    TDSQGNTVLHALVMISDNSAENIALVTSMYDGLLQAGARLCPTVQLEDIRNLQDLTPLKL
ratVR2      TDSLGNTVLHALVMIADNSPENSALVIHMYDGLLQMGARLCPTVQLEEISNHQGLTPLKL
            *** ***********:*** ** ***    ****** ***********:* * *.******

humanVR2    AAKEGKIEIFRHILQREFSG-LSHLSRKFTEWCYGPVRVSLYDLASVDSCEENSVLEIIA
ratVR2      AAKEGKIEIFRHILQREFSGPYQPLSRKFTEWCYGPVRVSLYDLSSVDSWEKNSVLEIIA
            *******************   .******************:***.*:*.*:********

humanVR2    FHCKSPHRHRMVVLEPLNKLLQAKWDLLIPKFFLNFLCNLIYMFIFTAVAYHQPTLKKQA
ratVR2      FHCKSPNRHRMVVLEPLNKLLQEKWDRLVSRFFFNFACYLVYMFIFTVVAYHQPSLDQPA
            ******:*************** *** *:.:**:** * *:******.******:*.: *

humanVR2    APHLKAEVGNSMLLTGHILILLGGIYLLVGQLWYFWRRHVFIWISFIDSYFEILFLFQAL
ratVR2      IPSSKATPGESMLLLGHILILLGGIYLLLGQLWYFWRRRLFIWISFMDSYFEILFLLQAL
             *   ** .*:**** *************:**********::******::******:***

humanVR2    LTVVSQVLCFLAIEWYLPLLVSALVLGWLNLLYYTRGFQHTGIYSVMIQKVILRDLLRFL
ratVR2      LTVLSQVLRFMETEWYLPLLVLSLVLGWLNLLYYTRGFQHTGIYSVMIQKVILRDLLRFL
            ***:**** *: ********* :*************************************

humanVR2    LIYLVFLFGFAVALVSLSQEAWRPEAPTGPNATESVQPMEGQEDEGNGAQYRGILEASLE
ratVR2      LVYLVFLFGFAVALVSLSREARSPKAPEDNNSTVTEQPTVGQEEEP--APYRSILDASLE
            *:****************:** *:** . *:* : ** ***:*    * **.**:****

humanVR2    LFKFTIGMGELAFQEQLHFRGMVLLLLLAYVLLTYILLLNMLIAIMSETVNSVATDSWSI
ratVR2      LFKFTIGMGELAFQEQLRFRGVVLLLLLAYVLLTYVLLLNMLIAIMSETVNHVADNSWSI
            *****************:***:*************:***************** :****

humanVR2    WKLQKAISVLEMENGYWWCR-KKQRAGVMLTVGTKPDGSPDERWCFRVEEVNWASWEQTL
ratVR2      WKLQKAISVLEMENGYWWCRRKKHREGRLLKVGTRGDGTPDERWCFRVEEVNWAAWEKTL
            *******************  **:* * ;*.***: **:****************:**;**

humanVR2    PTLCEDPSGAGVPRTLENPVLASPPKEDEDGASEENYVPVQLLQSN
ratVR2      PTLSEDPSGPGITGNKKNPTSK-PGK---NSASEEDHLPLQVLQSP
            ***.*****.*:.  . :**.   * *   :.****:::*:*:***
```

FIGURE 7

GAP of: ratvr2.pep   check: 0 from: 1 to: 554

ratVR2 Flrxb147g11

to: humanvr2.pep   check: 5746   from: 1 to: 764

humanVR2 Flh21e11

Symbol comparison table: /usr/local/gcg_9.1/gcgcore/data/rundata/blosum62.cmp
CompCheck: 6430

```
          Gap Weight:      12     Average Match:   2.912
       Length Weight:       4     Average Mismatch: -2.003

             Quality:   2182           Length:      766
               Ratio:   3.939            Gaps:        4
   Percent Similarity: 81.703    Percent Identity: 79.167
```

```
        Match display thresholds for the alignment(s):
                     | = IDENTITY
                     : =    2
                     . =    1
```

ratvr2.pep x humanvr2.pep

```
  1 ......STHASALSLAACTKQWDVVTYLLENPHQPASLEATDSLGNTVLH 44
        |||||||||||||.|||||||||||||:|||| ||||||
201 GQGTCFYFGELPLSLAACTKQWDVVSYLLENPHQPASLQATDSQGNTVLH 250

 45 ALVMIADNSPENSALVIHMYDGLLQMGARLCPTVQLEEISNHQGLTPLKL 94
    |||||.||| || ||| ||||||| ||||||.|||||:| | | ||||||
251 ALVMISDNSAENIALVTSMYDGLLQAGARLCPTVQLEDIRNLQDLTPLKL 300

 95 AAKEGKIEIFRHILQREFSGPYQPLSRKFTEWCYGPVRVSLYDLSSVDSW 144
    |||||||||||||||||||     |||||||||||||||||||.||||
301 AAKEGKIEIFRHILQREFSG.LSHLSRKFTEWCYGPVRVSLYDLASVDSC 349

145 EKNSVLEIIAFHCKSPNRHRMVVLEPLNKLLQEKWDRLVSRFFFNFACYL 194
    |.||||||||||||||.|||||||||||||||| |||.|: :|| || | |
350 EENSVLEIIAFHCKSPHRHRMVVLEPLNKLLQAKWDLLIPKFFLNFLCNL 399

195 VYMFIFTVVAYHQPSLDQPAIPSSKATFGESMLLLGHILILLGGIYLLLG 244
    :|||||| ||||||.|. . | | || | |||| ||||||||||||||.|
400 IYMFIFTAVAYHQPTLKKQAAPHLKAEVGNSMLLTGHILILLGGIYLLVG 449

245 QLMYFWRRRLFIWISFMDSYFEILFLLQALLTVLSQVLRFMETEWYLPLL 294
    ||||||||| .||||||.|||||||||| |||||||.|||| |: |||||||
450 QLMYFWRRHVFIWISFIDSYFEILFLFQALLTVVSQVLCFLAIEWYLPLL 499

295 VLSLVLGWLNLLYYTRGFQHTGIYSVMIQKVILRDLLRFLLVYLVFLFGF 344
    | .||||||||||||||||||||||||||||||||||||||||||:|||||||
500 VSALVLGWLNLLYYTRGFQHTGIYSVMIQKVILRDLLRFLLIYLVFLFGF 549

345 AVALVSLSREARSPKAPEDNNSTVTEQPTVGQEEE..PAPYRSILDASLE 392
    ||||||||.|| |.|| |.|. || |||:| | || ||:||||
550 AVALVSLSQEAHRPEAPTGPNATESVQPMEGQEDEGNGAQYRGILEASLE 599
```

FIGURE 8

```
393 LFKFTIGMGELAFQE LR      VLLLLLAYVLLTYVLLLNMLIALMSETV 44
    |||||||||||||||||      ||||||||||||||:||||||||||||||
600 LFKFTIGMGELAFQEQLHPRGMVLLLLLAYVLLTYILLLNMLIALMSETV 649

443 NHVADNSWSIWKLQKAISVLEMENGYWWCRRKKHREGRLLKVGTRGDGTP 492
    |  ||  ·||||||||||||||||||||||||  |||  |  |  |||:  ||·|
650 NSVATDSWSIWKLQKAISVLEMENGYWWC.RKKQRAGVMLTVGTKPDGSP 698

493 DERWCFRVEEVNWAAWEKTLPTLSEDPSGPGITGNKKNPT....SKPGKN 538
    |||||||||||||||·||·||||||  |||||  |:    ·||       |   ··
699 DERWCFRVEEVNWASWEQTLPTLCEDPSGAGVPRTLENPVLASPPKEDED 748

539 SASEEDHLPLQVLQSP 554
    ||||·:·|·|·|||
749 GASEENYVPVQLLQSN.764
```

FIGURE 8 (continued)

GAP of: humanvr1.seq   check: 4554   from: 1 to: 3909

humanVR1 Fbh18547pat - Import - complete

to: ratvr1.seq   check: 7921   from: 1 to: 2847

ratVR1.seq AF029310 in GenBank

Symbol comparison table:
/ddm_local/gcg/gcg_9.1/gcgcore/data/rundata/nwsgapdna.cmp
CompCheck: 8760

```
            Gap Weight:      50    Average Match: 10.000
         Length Weight:       3    Average Mismatch:  0.000

              Quality: 22717            Length:   3914
                Ratio:  7.979            Gaps:      10
   Percent Similarity: 82.125   Percent Identity: 82.125
```

```
        Match display thresholds for the alignment(s):
                      | = IDENTITY
                      : =    5
                      . =    1
```

humanvr1.seq x ratvr1.seq .

```
1001 CCAGGCCGTAGATGCTCCCCGCCGGTCAGTCACTTAGTCGTCAGATCGCC 1050
                                   ||      |        | ||
   1 ..............................CAGCTCCAAGGCACTTGCTCC 21

1051 CGTCCTGGTATCACAGTGCTTCTGTTCAGGTTGCACACTGGGCCACAGAG 1100
        |   ||| |  | |   | |       | ||||||| | ||||||||||||
  22 ATTTGGGGTGTGCCTGCACCT...AGCTGGTTGCAAATTGGGCCACAGAG 68

1101 GATCCAGCAAGGATGAAGAAATGGAGCAGCACAGACTTGGGGACAGCTGC 1150
     ||||  | |||||||| |  || ||   ||| |||||  | |     |||
  69 GATCTGGAAAGGATGGAACAACGGGCTAGCTTAGACTCAGAGGAGTCTGA 118

1151 GGACCCACTCCAAAAGGACACCTGCCCAGACCCCCTGGATGGAGACCCTA 1200
      | |!||| |||| || || ||||||  ||||| |  ||  ||||||||||
 119 GTCCCCACCCCAAGAGAACTCCTGCCTGGACCCTCCAGACAGAGACCCTA 168

1201 ACTCCAGGCCACCTCCAGCCAAGCCCCAGCTCCCCACGGCCAAGAGCCGC 1250
     ||| || |||||||||||| ||||||||| || ||| ||| ||| ||
 169 ACTGCAAGCCACCTCCAGTCAAGCCCCACATCTTCACTACCAGGAGTCGT 218

1251 ACCCGGCTCTTTGGGAAGGGTGACTCGGAGGAGGCTTTCCCGGTGGATTG 1300
     |||||||| |||||||||||||||||||||||||||||||| | || |||| ||
 219 ACCCGGCTTTTTGGGAAGGGTGACTCGGAGGAGGCCTCTCCCCTGGACTG 268

1301 CCCCCACGAGGAAGGTGAGTTGGACTCCTGCCCGACCATCACAGTCAGCC 1350
     |||  | ||||||||| | | |||   |||||||| | |||||| ||||||
 269 CCCTTATGAGGAAGGCGGGCTGGCTTCCTGCCCTATCATCACTGTCAGCT 318

1351 CTGTTATCACCATCCAGAGGCCAGGAGACGGCCCCACCGGTGCCAGGCTG 1400
     ||||| | || ||||||||||| || || || ||   || ||| |||||| |
 319 CTGTTCTAACTATCCAGAGGCCTGGGGATGGACCTGCCAGTGTCAGGCCG 368
```

<div align="center">

FIGURE 9

</div>

```
1401 CTGTCCCAGGACTCTGTCGCCGCCAGCACCGAGAAGACCCTCAGGCTCTA 1450
     |||||||||||| ||| |||| |   |||||| ||| ||||||||
 369 TCATCCCAGGACTCCGTCTCCGCTGG...TGAGAAGCCCCCGAGGCTCTA 415

1451 TGATCGCAGGAGTATCTTTGAAGCCGTTGCTCAGAATAACTGCCAGGATC 1500
     ||||||||||| |||||| || || || |||||||| ||||||||||||| |
 416 TGATCGCAGGAGCATCTTCGATGCTGTGGCTCAGAGTAACTGCCAGGAGC 465

1501 TGGAGAGCCTGCTGCTCTTCCTGCAGAAGAGCAAGAAGCACCTCACAGAC 1550
     ||||||||||||||| ||||||||||| ||||||||||| ||| || |||
 466 TGGAGAGCCTGCTGCCCTTCCTGCAGAGGAGCAAGAAGCGCCTGACTGAC 515

1551 AACGAGTTCAAAGACCCTGAGACAGGGAAGACCTGTCTGCTGAAAGCCAT 1600
     | ||||||||||||||| |||||||| |||||||||||||| ||||||||
 516 AGCGAGTTCAAAGACCCAGAGACAGGAAAGACCTGTCTGCTAAAAGCCAT 565

1601 GCTCAACCTGCACGACGGACAGAACACCACCATCCCCCTGCTCCTGGAGA 1650
     |||||| |||||| | || |||||   ||||||| | |||||||||||
 566 GCTCAATCTGCACAATGGGCAGAATGACACCATCGCTCTGCTCCTGGACG 615

1651 TCGCGCGGCAAACGGACAGCCTGAAGGAGCTTGTCAACGCCAGCTACACG 1700
     | || ||| | || ||||||||||| || ||||||| ||||||||||||
 616 TTGCCCGGAAGACAGACAGCCTGAAGCAGTTTGTCAATGCCAGCTACACA 665

1701 GACAGCTACTACAAGGGCCAGACAGCACTGCACATCGCCATCGAGAGACG 1750
     ||||||||||||||||||||||||||||||||||||| ||||| ||  | ||
 666 GACAGCTACTACAAGGGCCAGACAGCACTGCACATTGCCATTGAACGGCG 715

1751 CAACATGGCCCTGGTGACCCTCCTGGTGGAGAACGGAGCAGACGTCCAGG 1800
     |||||| | |||||||||||| ||||||||||| |||||||| |||||||
 716 GAACATGACGCTGGTGACCCTCTTGGTGGAGAATGGAGCAGATGTCCAGG 765

1801 CTGCGGCCCATGGGGACTTCTTTAAGAAAACCAAAGGGCGGCCTGGATTC 1850
     |||||||   | |||||||||| |||||||||||||||| |||||| |||
 766 CTGCGGCTAACGGGGACTTCTTCAAGAAAACCAAAGGGAGGCCTGGCTTC 815

1851 TACTTCGGTGAACTGCCCCTGTCCCTGGCCGCGTGCACCAACCAGCTGGG 1900
     ||||| ||||| ||||||||||||||||| |||||||||||||||||||||
 816 TACTTTGGTGAGCTGCCCCTGTCCCTGGCTGCGTGCACCAACCAGCTGGC 865

1901 CATCGTGAAGTTCCTGCTGCAGAACTCCTGGCAGACGGCCGACATCAGCG 1950
     ||| |||||||||||||||||||||||||||||||| | || ||||||||
 866 CATTGTGAAGTTCCTGCTGCAGAACTCCTGGCAGCCTGCAGACATCAGCG 915

1951 CCAGGGACTCGGTGGGCAACACGGTGCTGCACGCCCTGGTGGAGGTGGCC 2000
     || |||||||| ||||||||||||||||| || |||||||||||||||||
 916 CCCGGGACTCAGTGGGCAACACGGTGCTTCATGCCCTGGTGGAGGTGGCA 965

2001 GACAACACGGCCGACAACACGAAGTTTGTGACGAGCATGTACAATGAGAT 2050
     . || |||||| | |||||||| ||||| |||||| |||||||| |||||
 966 GATAACACAGTTGACAACACCAAGTTCGTGACAAGCATGTACAACGAGAT 1015

2051 TCTGATGCTGGGGGCCAAACTGCACCCGACGCTGAAGCTGGAGGAGCTCA 2100
     |||| ||||||||||||||||| ||||| |||||||||||||| ||| |||
1016 CTTGATCCTGGGGGCCAAACTCCACCCCACGCTGAAGCTGGAAGAGATCA 1065
```

FIGURE 9 (cont'd)

```
2101 CCAACAAGAAGGGAATGACGCCGCTGGCTCTGGCAGCTGGGACCGGGAAG 2150
     |||||| ||||||  |  |||||| ||||||||||| ||| | |  ||||||

1066 CCAACAGGAAGGGGCTCACGCCACTGGCTCTGGCTGCTAGCAGTGGGAAG 1115

2151 ATCGGGGTCTTGGCCTATATTCTCCAGCGGGAGATCCAGGAGCCCGAGTG 2200
     |||.|||||||||||||  ||||||||| |||||||||||| || ||||||||
1116 ATCGGGGTCTTGGCCTACATTCTCCAGAGGGAGATCCATGAACCCGAGTG 1165

2201 CAGGCACCTGTCCAGGAAGTTCACCGAGTGGGCCTACGGGCCCGTGCACT 2250
     | |  |||||| |||||||||||||||| |||||||||  ||||| ||||||||
1166 CCGACACCTATCCAGGAAGTTCACCGAATGGGCCTATGGGCCAGTGCACT 1215

2251 CCTCGCTGTACGACCTGTCCTGCATCGACACCTGCGAGAAGAACTCGGTG 2300
     |||| || ||  |||||||||||||| |||||||| || ||||||||||||||
1216 CCTCCCTTTATGACCTGTCCTGCATTGACACCTGTGAAAGAACTCGGTT 1265

2301 CTGGAGGTGATCGCCTACAGCAGCAGCGAGACCCCTAATCGCCACGACAT 2350
     ||||||||||||||| |||||||||||||·|| |||||.||||||| || || |||||
1266 CTGGAGGTGATCGCTTACAGCAGCAGTGAGACCCCTAACCGTCATGACAT⁼1315

2351 GCTCTTGGTGGAGCCGCTGAACCGACTCCTGCAGGACAAGTGGGACAGAT 2400
     |||  | |||||| || |||||||||||||| |||||·|||||||||||||||||||
1316 GCTTCTCGTGGAACCCTTGAACCGACTCCTACAGGACAAGTGGGACAGAT 1365

2401 TCGTCAAGCGCATCTTCTACTTCAACTTCCTGGTCTACTGCCTGTACATG 2450
     | ||||||||||||||||||||||||||||||||| | ||||||||| |||| |||
1366 TTGTCAAGCGCATCTTCTACTTCAACTTCTTCGTCTACTGCTTGTATATG 1415

2451 ATCATCTTCACCATGGCTGCCTACTACAGGCCCGTGGATGGCTTGCCTCC ·2500
     ||||||||||||| |||||||||||| |||| ||||| ||||||||| ||
1416 ATCATCTTCACCGCGGCTGCCTACTATCGGCCTGTGGAAGGCTTGCCCCC 1465

2501 CTTTAAGATGGAAAAAA...TTGGAGACTATTTCCGAGTTACTGGAGAGA 2547
     || |||| || ||||| |    |||| ||||||||||||||| |·| |||||||
1466 CTATAAGCTGAAAAACACCGTTGGGGACTATTTCCGAGTCACCGGAGAGA 1515

2548 TCCTGTCTGTGTTAGGAGGAGTCTACTTCTTTTTCCGAGGGATTCAGTAT 2597
     || |||||||||| ||||||||||||||||||| |||||||||||||·|||| |||
1516 TCTTGTCTGTGTCAGGAGGAGTCTACTTCTTCTTCCGAGGGATTCAATAT 1565

2598 TTCCTGCAGAGGCGGCCGTCGATGAAGACCCTGTTTGTGGACAGCTACAG 2647
     |||||||||||||| || ||  | |||||       ||||||||||||||||||||||
1566 TTCCTGCAGAGGCGACCATCCCTCAAGAGTTTGTTTGTGGACAGCTACAG 1615

2648 TGAGATGCTTTTCTTTCTGCAGTCACTGTTCATGCTGGCCACCGTGGTGC 2697
     |||||| |||||||||| | ||||| |||||||||||||| | |||||| |
1616 TGAGATACTTTTCTTTGTACAGTCGCTGTTCATGCTGGTGTCTGTGGTAC 1665

2698 TGTACTTCAGCCACCTCAAGGAGTATGTGGCTTCCATGGTATTCTCCCTG 2747
     ||||||||||||| | |||||||||||||||||||||||||||.||| ||||||||||
1666 TGTACTTCAGCCAACGCAAGGAGTATGTGGCTTCCATGGTGTTCTCCCTG 1715

2748 GCCTTGGGCTGGACCAACATGCTCTACTACACCCGCGGTTTCCAGCAGAT 2797
     ||| |||||||||||||||||||||||||||||||| ||||| || |||||||||||
1716 GCCATGGGCTGGACCAACATGCTCTACTATACCCGAGGATTCCAGCAGAT 1765
```

FIGURE 9 (cont'd)

```
2798 GGGCATCTATGCCGTCATGATAGAGAAGATGATCCTGAGAGACCTGTGCC 2847
     |||||||||||| ||||||||| |||||||||||||| |||||||||||||
1766 GGGCATCTATGCTGTCATGATTGAGAAGATGATCCTCAGAGACCTGTGCC 1815

2848 GTTTCATGTTTGTCTACATCGTCTTCTTGTTCGGGTTTTCCACAGCGGTG 2897
     | || ||||| |||||| |||| |||||||| || ||||||||||| |||
1816 GGTTTATGTTCGTCTACCTCGTGTTCTTGTTTGGATTTTCCACAGCTGTG 1865

2898 GTGACGCTGATTGAAGACGGGAAGAATGACTCCCTGCCGTCTGAGTCCAC 2947
     ||||| ||||||||.|| ||||||||| |||| ||||| ||||||||
1866 GTGACACTGATTGAGGATGGGAAGAATAACTCTCTGCCTATGGAGTCCAC 1915

2948 GTCGCACAGGTGGCGGGGGCCTGCCTGCAGGCCCCCCGATAGCTCCTACA 2997
      | |||| ||| |||||| |||||||||| |  || | || ||| ||||
1916 ACCACACAAGTGCCGGGGGTCTGCCTGCAAG...CCAGGTAACTCTTACA 1962

2998 ACAGCCTGTACTCCACCTGCCTGGAGCTGTTCAAGTTCACCATCGGCATG 3047
     |||||||||| ||||| || |||||||||||||||||||||||||||||||
1963 ACAGCCTGTATTCCACATGTCTGGAGCTGTTCAAGTTCACCATCGGCATG 2012

3048 GGCGACCTGGAGTTCACTGAGAACTATGACTTCAAGGCTGTCTTCATCAT 3097
     |||||||||||||||||||||||||| |||||||||||||||||||||||
2013 GGCGACCTGGAGTTCACTGAGAACTACGACTTCAAGGCTGTCTTCATCAT 2062

3098 CCTGCTGCTGGCCTATGTAATTCTCACCTACATCCTCCTGCTCAACATGC 3147
     |||| | ||||||||||| ||||||||||||||| ||||||||||||||
2063 CCTGTTACTGGCCTATGTGATTCTCACCTACATCCTTCTGCTCAACATGC 2112

3148 TCATCGCCCTCATGGGTGAGACTGTCAACAAGATCGCACAGGAGAGCAAG 3197
     |||| || ||||||||||||||| |||||||||| ||||| |||||||||
2113 TCATTGCTCTCATGGGTGAGACCGTCAACAAGATTGCACAAGAGAGCAAG 2162

3198 AACATCTGGAAGCTGCAGAGAGCCATCACCATCCTGGACACGGAGAAGAG 3247
     ||||||||||||||||||||||||||||||||||||||| || ||||||||
2163 AACATCTGGAAGCTGCAGAGAGCCATCACCATCCTGGATACAGAGAAGAG 2212

3248 CTTCCTTAAGTGCATGAGGAAGGCCTTCCGCTCAGGCAAGCTGCTGCAGG 3297
     |||||| |||||||||||||||||||||||||| |||||||||||||||||
2213 CTTCCTGAAGTGCATGAGGAAGGCCTTCCGCTCTGGCAAGCTGCTGCAGG 2262

3298 TGGGGTACACACCTGATGGCAAGGACGACTACCGGTGGTGCTTCAGGGTG 3347
     |||||| ||| ||||| ||||||||| |||||||||||||| |||||||||
2263 TGGGGTTCACTCCTGACGGCAAGGATGACTACCGGTGGTGTTTCAGGGTG 2312

3348 GACGAGGTGAACTGGACCACCTGGAACACCAACGTGGGCATCATCAACGA 3397
     |||||||| |||||||| |||||||||||||||| ||||| |||||||||
2313 GACGAGGTAAACTGGACTACCTGGAACACCAATGTGGGTATCATCAACGA 2362

3398 AGACCCGGGCAACTGTGAGGGCGTCAAGCGCACCCTGAGCTTCTCCCTGC 3447
     ||||| ||||||||||||||||||||||||||||||||||||||||||||
2363 GGACCCAGGCAACTGTGAGGGCGTCAAGCGCACCCTGAGCTTCTCCCTGA 2412

3448 GGTCAAGCAGAGTTTCAGGCAGACACTGGAAGAACTTTGCCCTGGTCCCC 3497
     ||||| || |||||||||| ||| |||||||||||||||||||||||| |||
2413 GGTCAGGCCGAGTTTCAGGGAGAAACTGGAAGAACTTTGCCCTGGTTCCC 2462

3498 CTTTTAAGAGAGGCAAGTGCTCGAGATAGGCAGTCTGCTCAGCCCGAGGA 3547
```

FIGURE 9 (cont'd)

```
              |||  |  || ||  |||||    |||||||||| ||   |   | ||||   || ||
2463 CTTCTGAGGGATGCAAGCACTCGAGATAGACATGCCACCCAGCAGGAAGA 2512

3548 AGTTTATCTGCGACAGTTTTCAGGGTCTCTGAAGCCAGAGGACGCTGAGG 3597
     ||||  |  |||   ||  |  |  |  || ||  || |||||||||||  |||||||
2513 AGTTCAACTGAAGCATTATACGGGATCCCTTAAGCCAGAGGATGCTGAGG 2562

3598 TCTTCAAGAGTCCTGCCGCTTCCGGGGAGAAGTGA.GGACGTCACGCAGA 3646
     |  ||||||  ||  |       |   ||||||||| |  |  ||||    |  ||||
2563 TTTTCAAGGATTCCATGGTCCCAGGGGAGAAATAATGGACACTATGCAGG 2612

3647 CAGCACTGTCAACACTGGGCCTTAGGAGACCCCGTTGCCACGGGGGGCTG 3696
      |  ||  ||                              |   |   |   ||  || |||
2613 GATCAATG.....................CGGGGTCTTTGGGTGGTCTG 2640

3697 CTGAGGGAACACCAGTGCTCTGTCAGCAGCCTGGCCTGGTCTGTGCCTGC 3746
     ||  |||||||  |       |       ||| |       | |   || |      |||||||||||
2641 CTTAGGGAAC.CAGCAGGGTTGACGTTATCTGGGTCCACTCTGTGCCTGC 2689

3747 CCA.GCATGTTCCCAAATCTGTGCTGGACAAGCTGTGGGAAGCGTTCTTG 3795
     |  |  |||   |||| |       ||   |   || |   |||||||||||  |          |
2690 CTAGGCACATTCCTAGGACTTCGGCGGGCCTGCTGTGGGAA.CTGGGAGG 2738

3796 GAAGCATGGGGAGTGATGTACATCCAACCGTCACTGTCCCCAAGTGAATC 3845
        |    |     |  |||||  ||||||||| | |    || ||||         |
2739 TGTGTGGGAATTGAGATGTGTATCCAACCATGA...TCTCCAAACATTTG 2785

3846 TCCTAACAGACTTTCAGGTTTTTACTCACTTTACTAAAAAAAAAAAAAAA 3895
        ||   |   ||||  || ||||  ||    |  || |||
2786 GCTTTCAACTCTTTATGGACTTTATTAAACAGAGTGAATGGCAAATCTCT 2835

3896 AGGGCGGCCGCTTA 3909
     |    |  |   |
2836 ACTTGGACACAT.. 2847
```

FIGURE 9 (cont'd)

GAP of: humanvr1.pep   check: 6877   from: 1 to: 839

humanVR1 _Fbh18547pat - fchrb87a6, 3909 bases, 4554 checksum.

to: ratvrl.pep   check: 5764   from: 1 to: 838

ratVR1 | AF029310 Rattus norvegicus vanilloid receptor subtype 1 mRNA, complete
cds.

Symbol comparison table:
/ddm_local/gcg/gcg_9.1/gcgcore/data/rundata/blosum62.cmp
CompCheck: 6430

        Gap Weight: 12   Average Match: 2.912
     Length Weight: 4 Average Mismatch: -2.003

          Quality: 3734   Length: 840
           Ratio: 4.456   Gaps: 3
Percent Similarity: 89.247 Percent Identity: 86.022

      Match display thresholds for the alignment(s):
                | = IDENTITY
                : =   2
                . =   1

humanvr1.pep x ratvr1.pep

```
  1 MKKWSSTDLGTAADPLQKDTCPDPLDGDPNSRPPPAKPQLPTAKSRTRLF 50
    |.. .| |    . | |...| || | ||| :||| || : | :||||||
  1 MEQRASLDSEESESPPQENSCLDPPDRDPNCKPPPVKPHIFTTRSRTRLF 50

 51 GKGDSEEAFPVDCPHEEGELDSCPTITVSPVITIQRPGDGPTGARLLSQD 100
    |||||||| |.|||:||| | ||| |||| |:|||||||||| |   |||
 51 GKGDSEEASPLDCPYEEGGLASCPIITVSSVLTIQRPGDGPASVRPSSQD 100

101 SVAASTEKTLRLYDRRSIFEAVAQNNCQDLESLLLFLQKSKKHLTDNEFK 150
    ||.| .||  |||||||||:||||.|||:||||| |||:||.|  |||.|||
101 SVSAG.EKPPRLYDRRSIFDAVAQSNCQELESLLPFLQRSKKRLTDSEFK 149

151 DPETGKTCLLKAMLNLHDGQNTTIPLLLEIARQTDSLKELVNASYTDSYY 200
    |||||||||||||||||.||| || |||::||.||||: ||||||||||
150 DPETGKTCLLKAMLNLHNGQNDTIALLLDVARKTDSLKQFVNASYTDSYY 199

201 KGQTALHIAIERRNMALVTLLVENGADVQAAAHGDFFKKTKGRPGFYFGE 250
    ||||||||||||||| ||||||||||||||||.|||||||||||||||||
200 KGQTALHIAIERRNMTLVTLLVENGADVQAAANGDFFKKTKGRPGFYFGE 249

251 LPLSLAACTNQLGIVKFLLQNSWQTADISARDSVGNTVHALVEVADNTA 300
    |||||||||||| |||||||||| ||||||||||||||||||||||||||
250 LPLSLAACTNQLAIVKFLLQNSWQPADISARDSVGNTVHALVEVADNTV 299

301 DNTKFVTSMYNEILMLGAKLHPTLKLEELTNKKGMTPLALAAGTGKIGVL 350
    ||||||||||||.|||||||||||||||:||:||:|||||||| .||||||
300 DNTKFVTSMYNEILILGAKLHPTLKLEEITNRKGLTPLALAASSGKIGVL 349

351 AYILQREIQEPECRHLSRKFTEWAYGPVHSSLYDLSCIDTCEKNSVLEVI 400
    ||||||| |||||||||||||||||||||||||||||||||||||||||||
350 AYILQREIHEPECRHLSRKFTEWAYGPVHSSLYDLSCIDTCEKNSVLEVI 399
```

<div align="center">

FIGURE 10

</div>

```
401 AYSSSETPNRHDMLLVEPLNRLLQDKWDRFVKRIFYFNFLVYCLYMIIFT 450
    |||||||||||||||||||||||||||||||||||||||| |||||||||
400 AYSSSETPNRHDMLLVEPLNRLLQDKWDRFVKRIFYFNFFVYCLYMIIFT 449

451 MAAYYRPVDGLPPFKMEK.IGDYFRVTGEILSVLGGVYFFFRGIQYFLQR 499
    |||||||:|||||:|:.  :|||||||||||| |||||||||||||||||
450 AAAYYRPVEGLPPYKLKNTVGDYFRVTGEILSVSGGVYFFFRGIQYFLQR 499

500 RPSMKTLFVDSYSEMLFFLQSLFMLATVVLYFSHLKEYVASMVFSLALGW 549
    |||:|.|||||||||.|||.||||| .||||| |||||||||||||:||
500 RPSLKSLFVDSYSEILFFVQSLFMLVSVVLYFSQRKEYVASMVFSLAMGW 549

550 TNMLYYTRGFQQMGIYAVMIEKMILRDLCRFMFVYIVFLFGFSTAVVTLI 599
    ||||||||||||||||||||||||||||||||||||||:||||||||||
550 TNMLYYTRGFQQMGIYAVMIEKMILRDLCRFMFVYLVFLFGFSTAVVTLI 599

600 EDGKNDSLPSESTSHRWRGPACRPPDSSYNSLYSTCLELFKFTIGMGDLE 649
    |||||.||| ||| |: || ||: | .||||||||||||||||||||||
600 EDGKNNSLPMESTPHKCRGSACK.PGNSYNSLYSTCLELFKFTIGMGDLE 648

650 FTENYDFKAVFIILLLAYVILTYILLLNMLIALMGETVNKIAQESKNIWK 699
    ||||||||||||||||||||||||||||||||||||||||||||||||||
649 FTENYDFKAVFIILLLAYVILTYILLLNMLIALMGETVNKIAQESKNIWK 698

700 LQRAITILDTEKSFLKCMRKAFRSGKLLQVGYTPDGKDDYRWCFRVDEVN 749
    ||||||||||||||||||||||||||||||:|||||||||||||||||||
699 LQRAITILDTEKSFLKCMRKAFRSGKLLQVGFTPDGKDDYRWCFRVDEVN 748

750 WTTWNTNVGIINEDPGNCEGVKRTLSFSLRSSRVSGRHWKNFALVPLLRE 799
    ||||||||||||||||||||||||||||||.|||||.|||||||||||:
749 WTTWNTNVGIINEDPGNCEGVKRTLSFSLRSGRVSGRNWKNFALVPLLRD 798

800 ASARDRQSAQPEEVYLRQFSGSLKPEDAEVFKSPAASGEK 839
    || ||| . | ||| |: :.|||||||||||||    |||
799 ASTRDRHATQQEEVQLKHYTGSLKPEDAEVFKDSMVPGEK 838
```

FIGURE 10 (cont'd)

```
CLUSTAL W (1.74) multiple sequence alignment


humanVR2.alt    ------------------------------------------------------------
humanVR2        MTSPSSSPVFRLETLDGGQEDGSEADRGKLDFGSGLPPMESQFQGEDRKFAPQIRVNLNY


humanVR2.alt    ------------------------------------------------------------
humanVR2        RKGTGASQPDPNRFDRDRLFNAVSRGVPEDLAGLPEYLSKTSKYLTDSEYTEGSTGKTCL


humanVR2.alt    ------------------------------------------------------------
humanVR2        MKAVLNLKDGVNACILPLLQIDRDSGNPQPLVNAQCTDDYYRGHSALHIAIEKRSLQCVK


humanVR2.alt    --------------GRFFQKGQGTCFYFGELPLSLAACTKQWDVVSYLLENPHQPASLQA
humanVR2        LLVENGANVHARACGRFFQKGQGTCFYFGELPLSLAACTKQWDVVSYLLENPHQPASLQA
                              ******************************************


humanVR2.alt    TDSQGNTVLHALVMISDNSAENIALVTSMYDGLLQAGARLCPTVQLEDIRNLQDLTPLKL
humanVR2        TDSQGNTVLHALVMISDNSAENIALVTSMYDGLLQAGARLCPTVQLEDIRNLQDLTPLKL
                ************************************************************


humanVR2.alt    AAKEGKIEIFRHILQREFSGLSHLSRKFTEWCYGPVRVSLYDLASVDSCEENSVLEIIAF
humanVR2        AAKEGKIEIFRHILQREFSGLSHLSRKFTEWCYGPVRVSLYDLASVDSCEENSVLEIIAF
                *************************************************************


humanVR2.alt    HCKSPHRHRMVVLEPLNKLLQAKWDLLIPKFFLNFLCNLIYMFIFTAVAYHQPTLKKQAA
humanVR2        HCKSPHRHRMVVLEPLNKLLQAKWDLLIPKFFLNFLCNLIYMFIFTAVAYHQPTLKKQAA
                ************************************************************


humanVR2.alt    PHLKAEVGNSMLLTGHILILLGGIYLLVGQLWYFWRRHVFIWISFIDSYFEILFLFQALL
humanVR2        PHLKAEVGNSMLLTGHILILLGGIYLLVGQLWYFWRRHVFIWISFIDSYFEILFLFQALL
                ************************************************************


humanVR2.alt    TVVSQVLCFLAIEWYLPLLVSALVLGWLNLLYYTRGFQHTGIYSVMIQ------------
humanVR2        TVVSQVLCFLAIEWYLPLLVSALVLGWLNLLYYTRGFQHTGIYSVMIQKVILRDLLRFLL
                ************************************************


humanVR2.alt    ------------------------------------------------------------
humanVR2        IYLVFLFGFAVALVSLSQEAWRPEAPTGPNATESVQPMEGQEDEGNGAQYRGILEASLEL


humanVR2.alt    ------------------------------------------------------------
humanVR2        FKFTIGMGELAFQEQLHFRGMVLLLLLAYVLLTYILLLNMLIALMSETVNSVATDSWSIW


humanVR2.alt    --KKAISVLEMENGYWWCRKKQRAGVMLTVGTKPDGSPDERWCFRVEEVNWASWEQTLPT
humanVR2        KLQKAISVLEMENGYWWCRKKQRAGVMLTVGTKPDGSPDERWCFRVEEVNWASWEQTLPT
                : **********************************************************


humanVR2.alt    LCEDPSGAGVPRTLENPVLASPPKEDEDGASEENYVPVQLLQSN
humanVR2        LCEDPSGAGVPRTLENPVLASPPKEDEDGASEENYVPVQLLQSN
                ********************************************
```

FIGURE 11

| 50 | 100 | 150 | 200 | 250 | 300 | 350 | 400 | 450 | 500 | 550 | 600 | 650 | 700 | 750 | 800 |

A — Alpha, Regions - Garnier-Robson

A — Alpha, Regions - Chou-Fasman

B — Beta, Regions - Garnier-Robson

B — Beta, Regions - Chou-Fasman

T — Turn, Regions - Garnier-Robson

T — Turn, Regions - Chou-Fasman

C — Coil, Regions - Garnier-Robson

Hydrophilicity Plot - Kyte-Doolittle

Hydrophobicity Plot - Hopp-Woods

Alpha, Amphipathic Regions - Eisenberg

Beta, Amphipathic Regions - Eisenberg

Flexible Regions - Karplus-Schulz

Antigenic Index - Jameson-Wolf

Surface Probability Plot - Emini

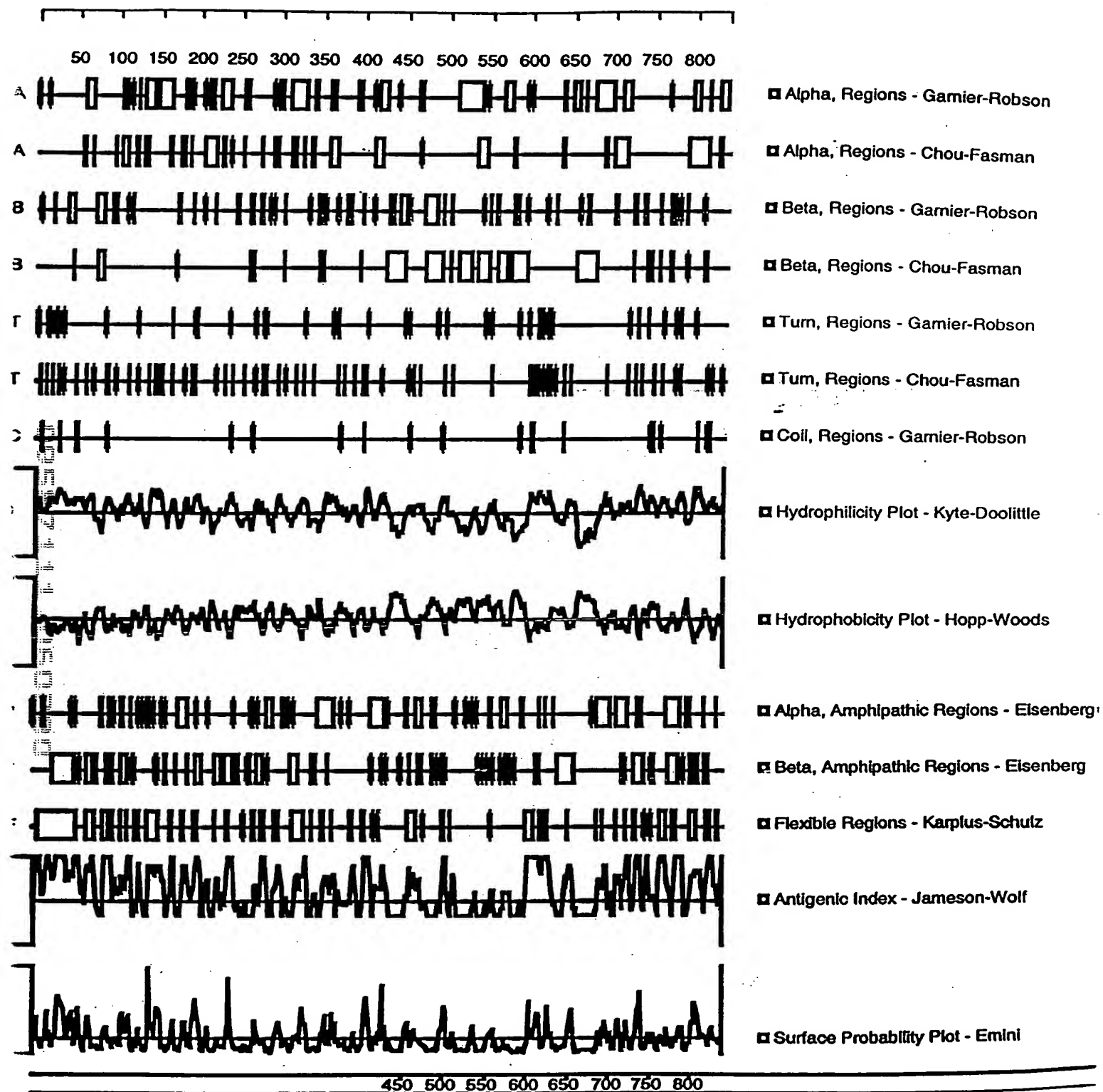450 500 550 600 650 700 750 800

FIGURE 12

Protein Family / Domain Matches, HMMer version 2

```
Searching for complete domains
hmmpfam - search a single seq against HMM database
HMMER 2.1.1 (Dec 1998)
Copyright (C) 1992-1998 Washington University School of Medicine
HMMER is freely distributed under the GNU General Public License (GPL).
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
HMM file:                 /prod/ddm/seqanal/PFAM/pfam4.2/Pfam
Sequence file:            /usr/ns-home/docs/seqanal/orfanal/oa-script.18670.seq
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
  Query:   hVR-1

Scores for sequence family classification (score includes all domains):
Model       Description                                    Score     E-value  N
--------    -----------                                    -----     ------- ---
ank         Ank repeat                                      51.5     1.9e-11   3

Parsed for domains:
Model     Domain    seq-f  seq-t    hmm-f hmm-t      score   E-value
--------  -------   ----- -----     ----- -----      -----   -------
ank         1/3      201   233 ..     1     33 []     34.4   2.6e-06
ank         2/3      248   283 ..     1     33 []     13.2        2
ank         3/3      333   361 ..     1     33 []      3.4       26

Alignments of top-scoring domains:
ank: domain 1 of 3, from 201 to 233: score 34.4, E = 2.6e-06
                   *->nGnTPLHlAarygnvevvklLLehGAdvnartk<-*
                      +G+T+LH+A +    n+ +v lL+e+GAdv a+


     hVR-1    201     KGQTALHIAIERRNMALVTLLVENGADVQAAAH       233

ank: domain 2 of 3, from 248 to 283: score 13.2, E = 2
                   *->nGnTPLHlAarygnvevvklLLe...hGAdvnartk<-*
                      G  PL lAa ++++ +vk+LL+++ + Ad+ ar+
     hVR-1    248     FGELPLSLAACTNQLGIVKFLLQnswQTADISARDS       283

ank: domain 3 of 3, from 333 to 361: score 3.4, E = 26
                   *->nGnTPLHlAarygnvevvklLLehGAdvnartk<-*
                      +G TPL lAa +g++ v ++ L+       ++
     hVR-1    333     KGMTPLALAAGTGKIGVLAYILQ----REIQEP       361
```
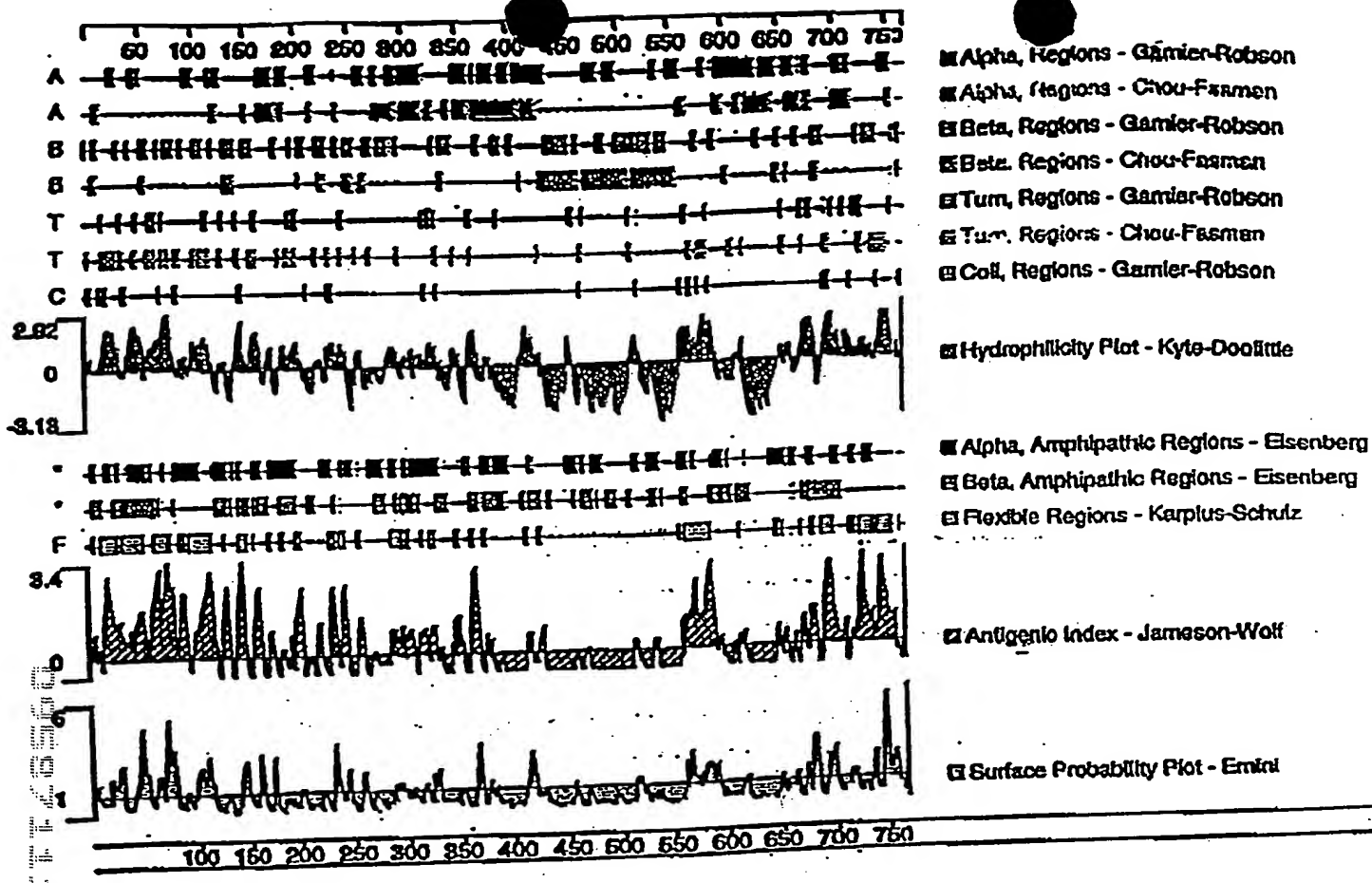
FIGURE 13

FIGURE 14

Protein Family / Domain Matches, HMMer version 2

```
Searching for complete domains
hmmpfam - search a single seq against HMM database
HMMER 2.1.1 (Dec 1998)
Copyright (C) 1992-1998 Washington University School of Medicine
HMMER is freely distributed under the GNU General Public License (GPL).
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
HMM file:              /prod/ddm/seqanal/PFAM/pfam3.3/Pfam
Sequence file:         /tmp/orfanal.579.aa
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -  - - -
Query:  Flh21e11

Scores for sequence family classification (score includes all domains):
Model     Description                                       Score    E-value  N
--------  -----------                                       -----    -------  ---
ank       PF00023 Ank repeat                                53.7     4e-12    3

Parsed for domains:
Model     Domain  seq-f  seq-t    hmm-f hmm-t      score   E-value
-------   ------  -----  -----    ----- -----      -----   -------
ank        1/3     162    194 ..      1    33 []    38.3   1.7e-07
ank        2/3     208    243 ..      1    33 []     6.4   4.3
ank        3/3     293    328 ..      1    33 []     8.8   2.1

Alignments of top-scoring domains:
ank: domain 1 of 3, from 162 to 194: score 38.3, E = 1.7e-07
                   *->nGnTPLHlAaryganvevvklLlehGAdvnartk<-*
                      +G+++LH+A ++ +++tvklLl+e+GA+v+ar
      Flh21e11   162  EGHSALHIAIERRSLQCVELLVENGANVHABAC     194

ank: domain 2 of 3, from 208 to 243: score 6.4, E = 4.3
                   *->nGnTPLHlAaryganvavvklLle...hGAdvnartk<-*
                      G  PL lAa + ++tvv +lLe+++t A+  a++
      Flh21e11   208  FGELPLSLAACTKQHDVVSYGLEqphQPASLQATDS    243

ank: domain 3 of 3, from 293 to 328: score 8.8, E = 2.1
                   *->nGnTPLHlAaryganvevvklLle...hGAdvnart<<-*
                      + +TPL lAa++g++e+ + L+++  G + +r
      Flh21e11   293  QDLTPLKLAAKEQKLEIFRHILQraCSGLSHLSRKF   328
```

FIGURE 15

MTSPSSSPVFRLETLDGGQEDGSLADR●DFGSGLPPMESQFQGEDRKFAPQ●NLNY
RKGTGASQPDPNRFDRDRLFNAVSRGV●LAGLPEYLSKTSKYLTDSEYTEGSTGKTCL
MKAVLNLKDGVNACILPLLQIDRDSGNPQPLVNAQCTDDYYRGHSALHIAIEKRSLQCVK
LVENGANVHARACGRFFQKGQGTCFYFGELPLSLAACTKQWDVVSYLLENPHQPASLQATDSQGNTVLHALVM
SDNSAENIALVTSMYDGLLQAGARLCPTVQLEDIRNLQDLTPLKLAAKEGKIEIFRHIL
REFSGLSHLSRKFTEWCYGPVRVSLYDLASVDSCEENSVLEIIAFHCKSPHRHRMVVLE
LNKLLQAKWDLLIPKFFLNFLCNLIYMFIFTAVAYHQPTLKKQAAPHLKAEVGNSMLLT
HILILLGGIYLLVGQLWYFWRRHVFIWISFIDSYFEILFLFQALLTVVSQVLCFLAIEW
LPLLVSALVLGWLNLLYYTRGFQHTGIYSVMIQKKAISVLEMENGYWWCRKKQRAGVML
VGTKPDGSPDERWCFRVEEVNWASWEQTLPTLCEDPSGAGVPRTLENPVLASPPKEDED
ASEENYVPVQLLQSN

FIGURE 16

CLUSTAL W (1.74) multip sequence alignment

```
humanVR2     MTSPSSSPVFRLETLDGGQEDGSEADRGKLDFGSGLPPMESQFQGEDRKFAPQIRVNLNY
hVR2.altFL   MTSPSSSPVFRLETLDGGQEDGSEADRGKLDFGSGLPPMESQFQGEDRKFAPQIRVNLNY
             ************************************************************

humanVR2     RKGTGASQPDPNRFDRDRLFNAVSRGVPEDLAGLPEYLSKTSKYLTDSEYTEGSTGKTCL
hVR2.altFL   RKGTGASQPDPNRFDRDRLFNAVSRGVPEDLAGLPEYLSKTSKYLTDSEYTEGSTGKTCL
             ************************************************************

humanVR2     MKAVLNLKDGVNACILPLLQIDRDSGNPQPLVNAQCTDDYYRGHSALHIAIEKRSLQCVK
hVR2.altFL   MKAVLNLKDGVNACILPLLQIDRDSGNPQPLVNAQCTDDYYRGHSALHIAIEKRSLQCVK
             ************************************************************

humanVR2     LLVENGANVHARACGRFFQKGQGTCFYFGELPLSLAACTKQWDVVSYLLENPHQPASLQA
hVR2.altFL   LLVENGANVHARACGRFFQKGQGTCFYFGELPLSLAACTKQWDVVSYLLENPHQPASLQA
             ************************************************************

humanVR2     TDSQGNTVLHALVMISDNSAENIALVTSMYDGLLQAGARLCPTVQLEDIRNLQDLTPLKL
hVR2.altFL   TDSQGNTVLHALVMISDNSAENIALVTSMYDGLLQAGARLCPTVQLEDIRNLQDLTPLKL
             ************************************************************

humanVR2     AAKEGKIEIFRHILQREFSGLSHLSRKFTEWCYGPVRVSLYDLASVDSCEENSVLEIIAF
hVR2.altFL   AAKEGKIEIFRHILQREFSGLSHLSRKFTEWCYGPVRVSLYDLASVDSCEENSVLEIIAF
             ************************************************************

humanVR2     HCKSPHRHRMVVLEPLNKLLQAKWDLLIPKFFLNFLCNLIYMFIFTAVAYHQPTLKKQAA
VR2.altFL    HCKSPHRHRMVVLEPLNKLLQAKWDLLIPKFFLNFLCNLIYMFIFTAVAYHQPTLKKQAA
             ************************************************************

humanVR2     PHLKAEVGNSMLLTGHILILLGGIYLLVGQLWYFWRRHVFIWISFIDSYFEILFLFQALL
VR2.altFL    PHLKAEVGNSMLLTGHILILLGGIYLLVGQLWYFWRRHVFIWISFIDSYFEILFLFQALL
             ************************************************************

humanVR2     TVVSQVLCFLAIEWYLPLLVSALVLGWLNLLYYTRGFQHTGIYSVMIQKVILRDLLRFLL
VR2.altFL    TVVSQVLCFLAIEWYLPLLVSALVLGWLNLLYYTRGFQHTGIYSVMIQK-----------
             *************************************************

humanVR2     IYLVFLFGFAVALVSLSQEAWRPEAPTGPNATESVQPMEGQEDEGNGAQYRGILEASLEL
VR2.altFL    ------------------------------------------------------------


humanVR2     FKFTIGMGELAFQEQLHFRGMVLLLLLAYVLLTYILLLNMLIALMSETVNSVATDSWSIW
VR2.altFL    ------------------------------------------------------------


humanVR2     KLQKAISVLEMENGYWWCRKKQRAGVMLTVGTKPDGSPDERWCFRVEEVNWASWEQTLPT
VR2.altFL    ---KAISVLEMENGYWWCRKKQRAGVMLTVGTKPDGSPDERWCFRVEEVNWASWEQTLPT
                ******************************************************

humanVR2     LCEDPSGAGVPRTLENPVLASPPKEDEDGASEENYVPVQLLQSN
VR2.altFL    LCEDPSGAGVPRTLENPVLASPPKEDEDGASEENYVPVQLLQSN
             *******************************************
```

FIGURE 17